Digital Business Methodology:

The Growth Code

- Consistency, Speed, & Scale (Enterprise)
- Business Engagement
- Business Ownership
- Data Quality

Foreword

Digital business can be done on its own without disrupting the current business. To drive this, Digital Business Methodology (DBM), is an outcome-driven, incremental approach to delivering business results with consistency, speed, and scale. It is business led in collaboration with key stakeholders from ideation to deployment to optimize outcomes with a focus on simplification of end-to-end workflows through a SSOT.

The DBM establishes strict governance in terms of engineering rigor, quality, security, compliance (auditable and traceable) ensuring higher productivity and predictability. It is enforced by a cloud based self-service digital business platform (DBP) that enables a simplified end-to-end collaborative environment independent from the existing complex, siloed environments.

ISO/IEC/IEEE 20148:2018 standardizes engineering requirements for systems and software throughout their life cycle. In the DBM framework, these standards are customized to refine product requirements and create a roadmap for achieving business outcomes. Applying these standards with the DBM enables universalization across various software systems, domains, technologies, and deployment types.

This book provides a comprehensive step-by-step guide to implementing a DBM that will accelerate growth and create greater agility. Regardless of digital maturity or organizational scale, the blueprint presented in this book outlines strategies for expediting your digital journey and deliver results. Furthermore, the inclusion of case studies demonstrates real-world application of the DBM and the outcomes it enabled.

Table of Contents

Introduction
Digital Business Methodology

Enterprises need to recognize the value of data and treat it as an essential asset. In doing so, businesses can actively engage and take ownership of compliance, security and ultimately, the outcomes that can drive exponential growth.

The potential benefits of effective data management include improved agility, greater responsiveness to trends, and enhanced management of risks and opportunities—all of which culminate in delivering outcomes with speed.

Currently, many enterprises lack a common repository for their core data assets with a single source of truth (SSOT), impeding the ability of business functions to effectively drive business engagement and ownership. It is important to address data quality across the enterprise, this is essential for the SSOT and ensures the effectiveness of ML and AI tools.

By establishing a digital business platform (DBP) and integrating data assets, enterprises can eliminate redundant information, rationalize technologies, simplify workflows, create end-to-end security, enforce governance that is auditable and traceable, and cultivate a unified and reliable enterprise SSOT.

Formidable obstacles companies typically face include managing their current complex, siloed technology environments along with disparate and duplicate data sources across the business. These complexities create barriers to seamless collaboration and integration for effective engagement and execution of the business strategy.

Furthermore, the pursuit of a "big bang" approach exacerbates the difficulty for businesses to assume full responsibility for effectively engaging and executing their strategic initiatives. This approach undermines their ability to take ownership of their operations, hindering progress and the realization of goals.

Increasingly, enterprises recognize the need to create a centralized DBP within the cloud, independent and away of their existing complex ecosystem. They are beginning to use new data technologies such as Snowflake and deploying them to construct robust data capabilities. However, the challenge lies in achieving a quick and seamless implementation of these tools.

Calibo complements and supplements Snowflake, significantly reducing the implementation timeframe by 50% or more and ensuring the creation of core data assets that are consistent and scalable.

The path to realize the benefits of the digital business methodology (DBM) is through a step-by-step engineering approach, adhering to IEEE standards. The

DBM focuses on an incremental approach specifically designed to deliver key outcomes. The business takes the ownership by engaging in these key steps:

1. Simplify the business outcome into clear, concise goals.
2. Create end-to-end workflows ensuring seamless integration.
3. Bring together relevant data and connect it from various sources, facilitating agile analysis and delivering key insights. This process is done at the sprint level, empowering the business to converge with a SSOT that becomes a core data asset in a Snowflake data mesh, leading to better decision making.
4. Harmonize diverse business functions to ensure outcomes for growth with greater consistency and scale.
5. Centralize master data management (MDM) and governance encompassing robust security, compliance, and elimination of duplicative data assets.

Enforcing this methodology requires a central business data platform that is fully automated. This fosters a collaborative self-service engineering environment that ensures enforcement of end-to-end security, compliance, and quality.

This platform also brings enterprise-wide innovation, operational rigor, and the ability to quickly launch new products for growth. Calibo's self-service platform offers a range of capabilities that not only fulfill essential business requirements, but also empower the business to take full ownership.

It also provides insight into security and compliance, with trusted data leveraging domain-level analytics and dashboards. The integration of machine learning (ML) with real time data identifies behavior patterns and insights that increase core-data intelligence.

By creating the appropriate algorithms through artificial intelligence (AI), a business can get a deeper understanding of its data. The benefit of the Calibo self-service platform is that it democratizes data at the domain-level, empowering users with analytics. It creates invaluable data assets.

The impact of deploying the DBM and DBP for core/domain data and analytics assets increases the adoption rate of Snowflake by 50% or more. Above all, it instills a pathway to unprecedented growth through data-driven assets and innovative thinking.

This ebook guides readers step by step through these concepts. By emphasizing consistency, business engagement and ownership, and boosting motivation, organizations can enjoy transformative business results.

This methodology has a proven track record and has delivered significant outcomes across enterprises of all sizes, as demonstrated through case studies for specific business outcomes, business domains, and enterprise core assets.

Chapter 1

Digital Business Outcomes for Growth, Engagement, and Ownership

Chapter Highlights

Establish a business case for business outcomes and growth using incremental approach.

Adopt an Agile culture increasing business and technology collaboration.

Take a bite-sized approach rather than a big bang approach.

Automate and simplify the end-to-end workflows.

Improve the experience of business users creating, managing and governing data assets.

Create a SSOT and core domain data assets in data mesh architecture.

A digital business is focused on outcomes and speed. It is designed to accelerate business growth through an Agile engineering culture. The purpose of the Digital Business Methodology (DBM) is to provide an approach to building a digital business that engages business, engineering, and data scientists together in a digital ecosystem in the cloud.

The business must be engaged in making the business case concise, simplifying end-to-end workflows, and taking ownership. Engineering must focus on building an agile culture and robust engineering practices. Data scientists must establish foundational data engineering principles that govern self-service and the adoption of artificial intelligence (AI) and machine learning (ML) in realizing new outcomes.

This chapter focuses on the role of the business in the DBM. There is an enormous opportunity for business to innovate and leverage data resources. For the business to grow, senior leaders must take ownership of data and provide self-service capabilities at a domain level.

Figure 1 below outlines the flow of the DBM, and highlights the role played by the business in staying engaged, owning data, and converging into a data platform like the Snowflake Data Mesh.

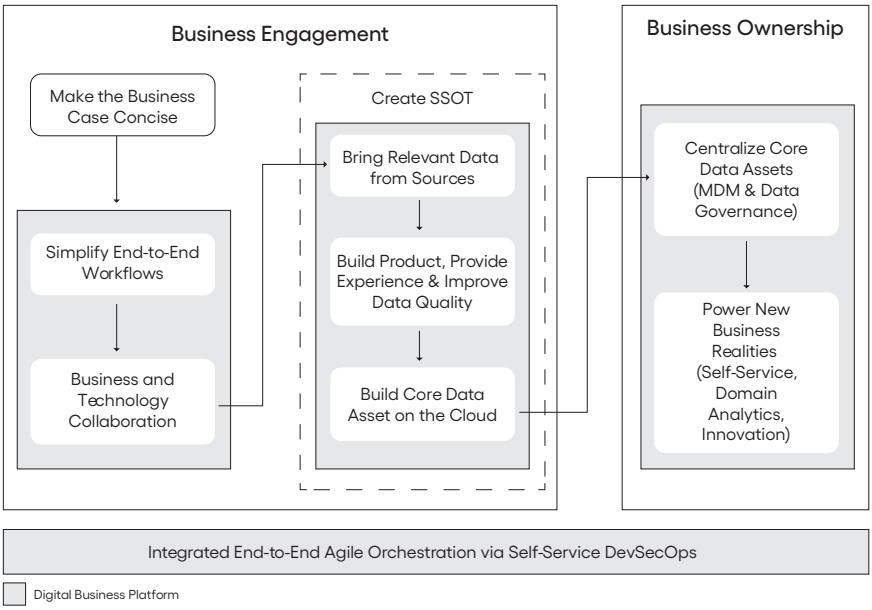
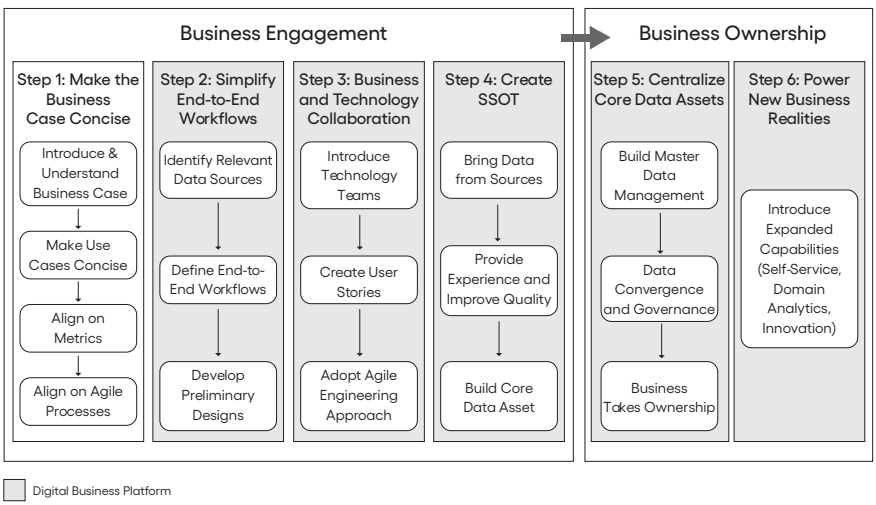


Figure 1: Digital Business Methodology

Figure 2 below offers further elaboration of the steps in the DBM, and these are further detailed in the rest of this chapter.

Figure 2: Digital Business Methodology—Business Engagement and Business Ownership Process Steps



Business Engagement

1. A business engagement will typically employ the following steps:
2. Make the business case concise.
3. Simplify the end-to-end workflows.
4. Business and technology collaboration.
5. Create the SSOT.

> "There is an enormous opportunity for business to grow, by engaging and taking ownership of data resources."

Step 1: Make the Business Case Concise

Business stakeholders are comfortable articulating business outcomes and what they wish to achieve in the form of a business case. Technology teams, on the other hand, seek elaborate solutions to meet business needs. To align the business and technology stakeholders, it is critical to make the business case concise.

First, introduce and understand the business case. The business leader must assemble the key stakeholders into a workshop where the business case is articulated, and common challenges are openly discussed.

By the conclusion of the workshop, all stakeholders must have a shared understanding of the business case and be aligned on the scope. All stakeholders must articulate what the business case covers and there must be consensus among the participants.

Next, make the use cases concise. A cross-functional team must now take the articulation of the business case to the next level of detail.

Once the use cases are completed, each stakeholder should have some sense of what dimension of the business case they can address.

Higher quality and granularity of the use case typically translates into a higher likelihood of success for the digital initiative. The use cases need to be understood by the engineering teams so that they can determine the solution approach as part of an Agile development framework.

Stakeholders should collaborate on a preliminary view of how to group the use cases into smaller bundles as part of a bite-sized approach.

Following this comes alignment on metrics. It is the responsibility of the business stakeholders to articulate the comprehensive list of metrics based on what is

known at the time. This will help the stakeholders understand what the business is seeking to influence and measure for a use case. Metrics must be assigned at the use case level and give stakeholders an understanding of how a solution will be evaluated. There should also be a dialog around understanding the relevant data.

Agile processes also must be aligned. It is imperative to gain alignment on the process that would be used to support this effort. The Agile development process is often used for this purpose because it provides agility and speed in helping to deliver bite-sized increments to the business. This process also helps the business gain an understanding of what the final product will look like early in the process.

Step 2: Simplify the End-to-End Workflows

Once the use cases have been finalized and are well understood, the next step is to create and simplify the series of workflows that correspond to what has been defined in the use cases.

Identify relevant data sources, which typically involves identifying the source of data as the starting point. This serves to connect the data source to the use case.

Next, define end-to-end workflows. Workflows must describe a variety of scenarios, including user actions and decisions made as part of a process. They also outline how data moves through a system and might even define how data is cleansed and harmonized.

Workflows should provide greater detail of the "what" from the use case.

The "how" will be defined at a later stage as part of the engineering solution approach. Like the use case creation process, the workflow definition process might require several iterations to distill the workflows into a form that is as simple as possible.

Definition of end-to-end workflows is the starting point for using a digital business platform. It helps to harmonize the DBM steps and facilitates collaboration among all stakeholders.

Once the workflow definition effort is completed, it's time to engineer preliminary designs and get feedback. It's critical to further simplify the underlying processes via UX design and proof of value channels by showcasing through a wireframe and/or prototype. This allows end users to provide feedback and decide what to maintain, alter or eliminate as part of the overall process.

Step 3: Business and Technology Collaboration

Business and technology collaboration starts with the creation of user stories and adopting an Agile approach as the collaboration vehicle between stakeholders.

Introduce a technology team (for example, an Agile scrum team), which will convert inputs into user stories that they can be consumed at the sprint level. User stories help understand what needs to be achieved and how users will leverage the functionality. They also allow the scrum team to organize their efforts based on a two-week sprint.

The data should also be clearly defined as a collaborative effort supported by the business and technology teams. Engineers might contribute to the story backlog by providing input related to data interoperability, integration, security, and compliance, if needed.

There might also be user stories related to identifying and injecting data from various sources that would enable the desired functionality.

Adopt an Agile engineering approach at the sprint level. User stories should now be organized into a prioritized backlog and loaded into a sprint as part of a process called sprint planning. Engineering should then be able to define a finite workload to be completed within a defined period.

This will also inform on the number of sprints needed to deliver the scope. Application and data engineers can now start the process of building the solution by adopting Agile principles at the sprint level.

Step 4: Create the SSOT

Once the sprint is completed, the engineering teams must deliver the experience conveyed in the user story. This will enable these teams to request feedback from stakeholders and identify areas of improvement. Once approved, the core data asset can be created in a cloud data mesh leveraging Snowflake.

Data sources must be identified while considering interoperability and integration of these data sets into a centralized system. In many cases, the data will need to be injected from the relevant source systems to support the solution for the use case.

Engineering teams can now help business stakeholders understand what functionality can be delivered. This will determine firsthand the value of what is being introduced, build the business trust and confidence in the solution, and secure buy-in from business stakeholders to the solution.

Providing an early example of the experience is an important step in securing feedback for improving the experience and making it more intuitive. It's important to view what is delivered as a preliminary indication of what the end state could be, to ensure that all stakeholders are aligned and on the right track.

Stakeholder feedback should also address data issues and identify new business cases related to improving data quality. Technology teams can use the business

stakeholder's input to determine whether the data is meeting business needs sufficiently.

Once the business confirms that the right data is being presented, this indicates that the goal of achieving SSOT has been met.

Next, build the core data asset. A cloud SSOT implies that everyone in the organization uses the same data when making business decisions.

While the definition is simple, organizations have struggled to bring together data due to multiple challenges such as the use of discrete systems, proprietary processes, technical debt, and competing business outcomes.

After the SSOT has been established, focus activity on aggregating all the data sets into a technology solution such as Snowflake mesh, which will serve as focal point for all SSOT activities.

To enable the evaluation of frequent releases, it might be appropriate to create a dedicated environment where users can experience each incremental release. This has the added benefit of creating a feedback loop in which the business can evaluate what has been delivered and introduce incremental changes along the way.

Business Ownership

Once technology teams begin releasing solutions to a dedicated environment in an incremental manner, it is time for business to take ownership of data and drive new possibilities.

> "To align the business and technology stakeholders, it is critical to make the business case concise."

Business ownership is defined as a series of steps where the business is using and integrating new solutions into day-to-day activities to drive value more quickly. What really accelerates this transition into business ownership is the introduction of new capabilities focused on self-service, since this significantly minimizes the dependency on engineering to help facilitate the generation of these insights.

Business ownership focuses on the following steps:

1. Centralization of core data assets.
2. Powering new business realities through expanded capabilities.

Step 5: Centralization of Core Data Assets

Data centralization should maximize the value of the data across the organization in a manner that is consumable by all stakeholders. It must avoid the limitations that come from having disparate siloed solutions throughout the organization.

Building mechanisms such as a master data management (MDM) repository help overcome limitations associated with siloed solutions and disparate integrations. Business teams need to define core master and reference data elements that need to be tracked and maintained in the system. And they also need to seek alignment of terminologies across the enterprise to build a robust MDM.

The individual pieces driving the outcomes should converge toward an organization wide SSOT. While this is being incrementally built based on a series of bite-sized outcomes, users should see an improved experience that results in faster value with each increment.

Here are some key areas of focus for SSOT workflows that are aligned with standards defined by established governing bodies:

- Identify relevant data sources for both the short-term and long-term and build data structures that anticipate future business needs.
- Define how the data will be brought to source systems while ensuring alignment to customer journeys and the needs of various stakeholders.
- Emphasize centralization and consistency that standardizes a uniform definition and source for master data and propagation across the organization.
- Establish uniform data stewardship and controls across the organization.
- Analyze the data quality and define approaches to cleanse data and provide it in a consumable format for consumers.
- Define security and compliance needs based on domain and business function.
- Establish unified monitoring and reporting for data as well as robust traceability.
- Define data retention and usage policies based on business needs.

By aggregating these activities into a Snowflake or similar solution, the core data serves as an enterprise-level SSOT. Centralization in and of itself has immediate value and further incentivizes the business to take ownership and drive even greater collaboration.

The engineering teams should work to ensure that the business is comfortable with managing the solution with the intention of eventually transitioning full ownership of the solution to the business. As the business starts performing its own analysis and data science more independently through the introduction of self-service tools, it begins taking more ownership at the domain level.

Step 6: Powering New Business Realities

The foundational elements now come together in a simplified form that can be used to power user experience and insights and greater data discovery. This enables users to think more comprehensively about use cases driven by AI/ML that automate and enhance outcomes.

New and enhanced capabilities that can be delivered through SSOT include solidifying security, quality, and compliance, which leads to discovery and domain level analytics; improved self-service; data dashboards that serve as a steppingstone to more advanced analytics; patterns and behaviors with real time analytics; AI/ML adoption and higher accuracy of prediction; and reusable visualization components.

Business engagement and business ownership act as two pillars to amplify and accelerate business outcomes.

Case Study—Accelerating Supply Chain Outcomes for a Specialty Chocolatier

Overview of the Company

A $1 billion specialty chocolatier with manufacturing facilities across Turkey, the United States and Europe sold gourmet chocolates to retail outlets via the CPG business and to consumers via the direct-to-Consumer model.

Targeted Business Outcomes

- Address supply chain issues around inventory availability and inventory visibility.
- Reduce excess inventory of slow-moving SKUs.
- Improve availability of fast-moving SKUs.
- Improve inventory visibility to sales teams.
- Improve demand forecasting.

Execution Steps

Business Engagement

Step 1: Make the Business Case Concise

The CEO mandated that the focus must be on improving inventory visibility and availability. Challenges included extensive use of Excel worksheets and an existing operating model in sales, purchasing and manufacturing.

A cross-functional team took responsibility for creating detailed use cases. Sales teams had to move away from Excel and order based on real-time inventory information. Visibility into inventory shelf life had to be provided to all teams. Manufacturing had to place orders for raw materials well in advance of production start dates.

Metrics were defined for each of these use cases, and the team aligned on an Agile development process.

Step 2: Simplification of End-to-End Workflows

Relevant data sources from supply chain and manufacturing applications were identified. The team created detailed workflows and engineered preliminary designs of wireframes and prototypes.

Step 3: Business and Technology Collaboration

Data architects and software engineers were introduced into the mix. The team created user stories at the sprint level. An effort of six sprints was estimated as part of the Agile engineering approach.

Step 4: Build the SSOT

The engineering team brought data from source systems and created an example, received feedback, and adjusted as needed. It started building the end-to-end solution. A core asset was created for supply chain and manufacturing domains.

Business Ownership

Step 1: Centralization of the Core Data Asset

The team evolved the solution and created a master data management framework. Data was gradually converged into the framework and common governance mechanisms were identified. Business now took ownership of the data assets.

Step 2: Powering New Business Realities

Business was able to leverage expanded capabilities such as AI/ML and perform more sophisticated supply chain and inventory analytics.

Outcomes Achieved

- Improved inventory availability and inventory visibility.
- Use of Excel sheets was eliminated.
- Tracking inventory position by shelf-life reduced obsolete inventory by 80%.
- Real-time information to teams improved cycle times by 25%.
- Trajectory of revenue uplift tracked toward 5%.

Case Study—Accelerating Consumer and Business Insights for a Leading CPG Brand

Overview of the Company

A leading food brand in Australia had its bakery operations spread across multiple countries in the Asia Pacific region. It sold its biscuits and cookies through mass merchandisers and retailers.

Targeted Business Outcomes

- Provide predictive insights on consumer buying patterns and consumer behavior.
- Establish understanding on competitor strategies.
- Engineer business insights on revenue, profitability, and cost.

Execution Steps

Business Engagement

Step 1: Make the Business Case Concise

The management team initiated a workshop to introduce the business case and create a shared understanding. Challenges included a lack of shopper segments; poor understanding of consumer flavor preferences; need to understand product substitution by shoppers; determining competitor strategies; and limited insights on revenue, profitability and costs at a brand and category level.

A cross-functional team took responsibility for creating detailed use cases, and alignment was secured on metrics to measure use cases. The team aligned on an Agile development process.

Step 2: Simplification of End-to-End Workflows

Relevant data sources from marketing and sales applications were identified and detailed workflows were created. The team engineered preliminary designs of wireframes and prototypes.

Step 3: Business and Technology Collaboration

Data architects and software engineers were introduced into the mix. The team created user stories at the sprint level, and an effort of eight sprints was estimated as part of the Agile engineering approach.

Step 4: Build the SSOT

The engineering team brought data from source systems, created an example, and received feedback and adjusted as needed. It started building the end-to-end solution. A core asset was created for marketing and sales domains.

Business Ownership

Step 1: Centralization of the Core Data Asset

The team evolved the solution and created a master data management framework. Data was gradually converged into the framework and common governance mechanisms were identified. Business took ownership of the data assets.

Step 2: Powering New Business Realities

Business was able to leverage expanded capabilities such as AI/ML and perform sophisticated consumer and business insights.

Outcomes Achieved

- The company built a successful consumer clustering algorithm to gain a richer understanding of consumers and buying behavior. Seven new clusters were identified.
- Consumer clustering in turn informed consumer flavor and product substitution preferences.
- The company was able to successfully corelate competitor pricing strategies on product sales.
- The firm engineered detailed business insights on revenue, profitability, and cost at a category level.

Case Study: Accelerating Financial Reporting Outcomes for a Global Automotive Supplier

Overview of the Company

A leading global automotive supplier has operations in more than 50 countries and across 150 legal entities. It faced elongated book close and consolidation cycle time of about four work weeks, and sub-optimal use of finance teams due to the manual nature of collating data and generating financial information.

Targeted Business Outcomes

- Reduction in book close and consolidation of cycle time by 30%.
- Operational efficiency in the finance function by 20%.
- Standardized local and global view of financial reporting.

Execution Steps

Business Engagement

Step 1: Make the Business Case Concise

The CFO initiated a workshop to introduce the business case. Challenges included a book close time of about four work weeks; a manual approach to collating data; variation in accounting policies and procedures; more than 20 ERP applications supporting the finance process.

A cross-functional team took responsibility for creating detailed use cases. Stakeholders aligned on metrics to measure use cases, and the team aligned on an Agile development process.

Step 2: Simplification of End-to-End Workflows

Relevant data sources for the use cases from 20-plus ERP applications were identified. The team established detailed workflows to build a unified finance ledger data model and engineered preliminary designs of wireframes and prototypes.

Step 3: Business and Technology Collaboration

Data architects and software engineers were introduced into the mix. The team created user stories at the sprint level, and an effort of 10 sprints was estimated as part of the Agile engineering approach.

Step 4: Build the SSOT

The engineering team built pipelines to obtain data from source systems. Transaction data was enriched and harmonized. The team started building the end-to-end solution. A core asset was created for the finance domain.

Business Ownership

Step 1: Centralization of the Core Data Asset

The team evolved the solution and created an MDM framework. Data was gradually converged into the framework with governance mechanisms established. Business took ownership of the data assets.

Step 2: Powering New Business Realities

Business was able to leverage expanded capabilities such as AI/ML. Sophisticated finance analytics and what-if modeling were now a reality.

Outcomes Achieved

- 50% reduction in book close cycle time.
- 30% improvement in resource productivity.
- 80% automation in consolidation of financial transactions across the enterprise.
- 100% data traceability and auditability.

Case Study: Enhancing Forecast Accuracy Using Data Science for a Global Agro Firm

Overview of the Company

A global Agro company provides farm produce to some of the world's largest retailers. The yield from its greenhouses was forecast using historical information and dated rules. It sought to improve its yield forecasting accuracy across the harvesting cycle for its green houses.

Targeted Business Outcomes:

- Improve yield forecasting accuracy to 90%.
- Optimize the demand management approach.

Execution Steps

Business Engagement

Step 1: Make the Business Case Concise

The business leader initiated a workshop to introduce the business case. Challenges included yield forecasting accuracy of under 80%; no formal mechanism to review and diagnose the forecast variance; and no visibility into the patterns present in the raw data.

A cross-functional team took responsibility for creating detailed use cases. Stakeholders aligned on metrics to measure use cases. The team aligned on an Agile development process.

Step 2: Simplification of End-to-End Workflows

Relevant data sources for the use cases were identified and detailed workflows established. The team engineered preliminary designs of wireframes and prototypes.

Step 3: Business and Technology Collaboration

Data scientists, data architects and software engineers were introduced into the mix. Data scientists performed a deep analysis of historical data, reviewed data patterns, and developed an approach to build a yield forecasting engine.

User stories were created at the sprint level and an effort of 12 sprints was estimated.

Step 4: Build the SSOT

The engineering team built pipelines to obtain data from source systems. A core asset was created for the forecasting domain.

Business Ownership

Step 1: Centralization of the Core Data Asset

The team evolved the solution and resolved master data references. Business took ownership of the data assets.

Step 2: Powering New Business Realities

Business was able to harness expanded capabilities such as AI/ML. Growers' experience was leveraged to improve the forecasting model.

Outcomes Achieved

- One-week yield forecasting accuracy increased to 92%.
- Six-week yield forecasting accuracy increased to 86%.

Chapter 2

Data Democratization

Chapter Highlights

- Business engagement in core data and data quality on Snowflake based data mesh.
- Business ownership in MDM and governance and core data with no duplication.
- End-to-end Agile engineering environment; integrated and automated for faster development and adoption.
- Self-service capabilities for business domain, analytics.

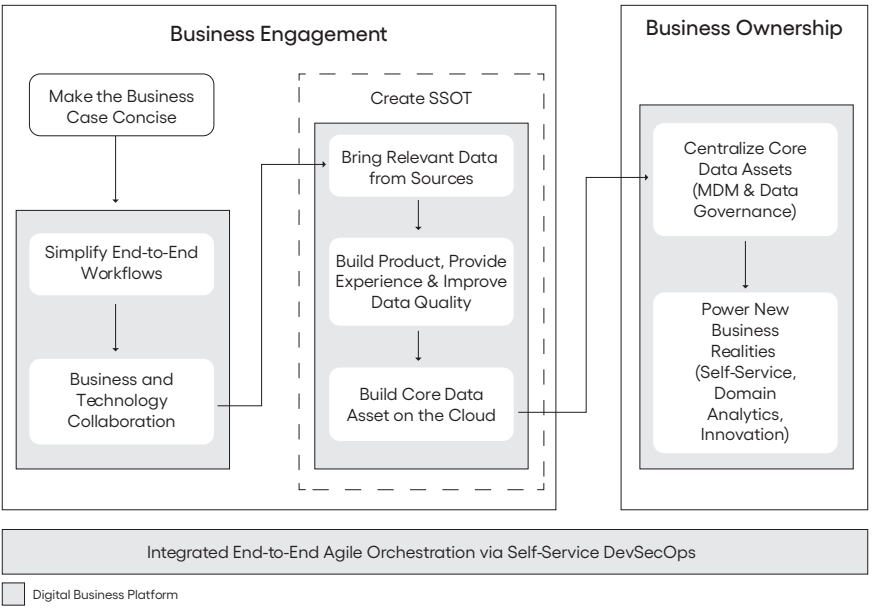
The previous chapter focused on the role of the business in the Digital Business Methodology to enable acceleration of outcomes through data. This chapter focuses on data democratization and the engineering of the data core levering a data mesh architecture pattern.

The data engineering methodology is created with focus on the following key guiding principles:

Data/Product Engineering

As organizations have increased their momentum for the digital journey, they are relying on data from various sources to be accessed by personas and products. Business leaders and users need to trust data, so it needs to be harmonized and standardized across the organization to provide the desired consistency and reliability.

Thus, there is an imperative to create a SSOT and maintain it to serve various personas and products.



> "Do not go overboard and launch into large multi-year data programs by derailing the business needs. Focus on immediate business need and you will get there incrementally."

Building an efficient SSOT can be a daunting task. In a typical organization, business relies on multiple transactional systems (ERP, CRM, MES, HCM, Planning). Integrating data from all these systems requires a crisp strategy and governance.

A data platform enabling an outcome driven, bite-size implementation approach has helped many organizations deliver a highly trusted and ever expanding SSOT.

The illustration below highlights the main guidelines to build an SSOT with continuous business engagement for easier adoption, leveraging a data platform like Snowflake with data mesh architecture.

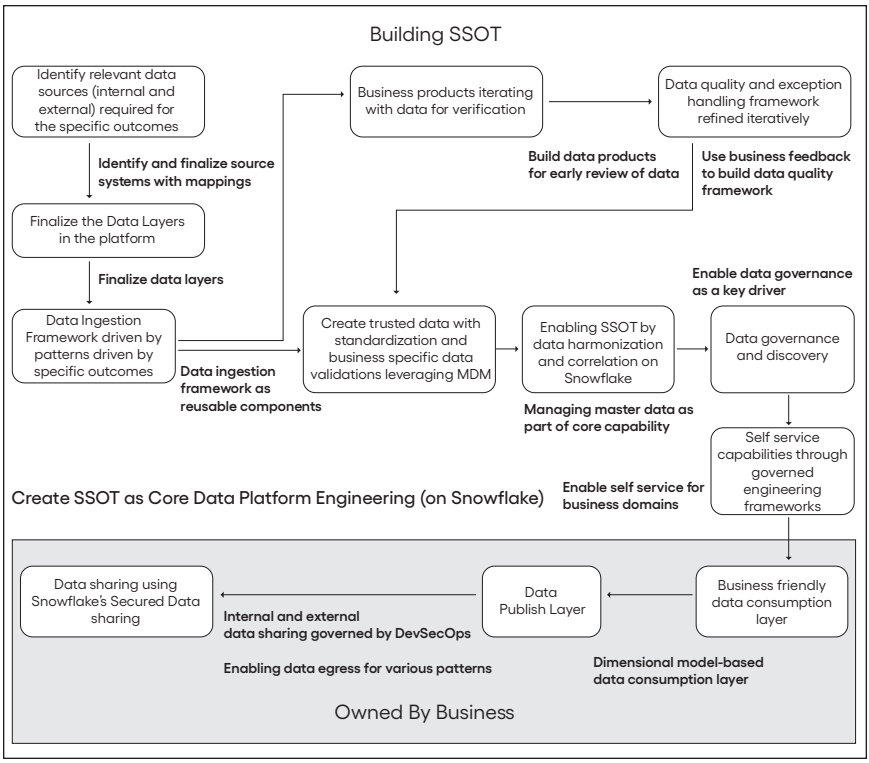


Figure 2. SSOT – Engineering Methodology

Business-Led Program Definition

The success of an SSOT solution relies on business engagement during the lifecycle. While the business defines the context and use cases to be enabled by the SSOT platform, the program alignment focuses on elaborating this in detail, which will typically uses the following step:

Step 1: Identification and Finalization of the Source Systems

Without an end goal in mind and a business outcome as a part of the pilot implementation strategy, the entire effort can turn into a data swamp. Different data sources will reside in a common repository, but business users will be unable to properly consume the data. Focus on the bite-size approach for building the SSOT, and select the required source systems to support business scenarios.

Business shares the metrics and work with engineers to identify the relevant

source systems and create the mapping of the fields/attributes from the source system entities to the metrics.

Building the Core of the SSOT

The core of the SSOT solution will primarily comprise of data collection, data quality and data layers.

Step 1: Finalizing the Designated Data Layers

Snowflake's data storage and physical modelling abilities provides a good basis for layering data zones. Data mesh principles can be applied while finalizing the data layers.

Typically, three or more layers are created. The landing layer is used for storing data, as it is from the source system. Data is stored in the same format as received from the source. The next layer is the curated layer, predominantly used to standardize data and to enforce quality checks. The gold layer creates data sets in a dimensional model. It hosts cleaned and verified data, and can be created for different business domains with specific data transformation requirements.

An additional consumption layer can be created for making the data available to business users. Snowflake's secured data sharing can be leveraged for creating business domain specific consumption layers.

Step 2: Building a Data Ingestion Framework as a Reusable Component

As the SSOT platform scales up to accommodate new source systems and new use cases with different metrics, it is important to take scalability and expandability into account in the design of the platform. The data ingestion framework enables this.

Adopt pattern-based development to create reusable components. First, analyse the connection and interface for the source system datasets and determine whether the data movement happens as a push from source system, or a pull initiated by the SSOT platform. Then determine the frequency of the data extract for each source system and analyse the availability of the incremental data. Finally, identify encryption/obfuscation for in-transit data.

Build parameter-based components focusing on required connection settings for each pattern (e.g. for FTP, parameters for server details, port number) using DevOps and CI/CD. Create a data reconciliation process as a component. It should include completeness of the data load as well as data types. Engage with the business team to validate the SLAs and data quality.

Step 3: Build a Business Product for Early Review of Data

An early access to data for the business team improves the quality and trust in data. An iterative cycle allows the team to validate the product by accessing the data and providing guidance for driving cleaner, data. The continuous feedback helps to create the data quality framework.

Step 4: Define the Data Quality Framework

Data quality is another reusable component. Different data quality aspects are applied to different data layers. Business provides the business logic for quality checks and engineers implement the rules as reusable components, where applicable.

Enable data profiling on data at the landing layer and define the standardization rules for a curated layer based on the data domains (for example, customer, address, phone, email etc.). Push business domain specific quality checks for layers that transform data for metrics calculations. Enable rules-based quality checks for faster integration with data pipelines and better scalability.

As the core is being designed and built, a parallel track owned by business and enabled by engineers for data governance and MDM solution is initiated.

Step 5: Managing Master Data as a Core Capability

An efficient SSOT requires master and reference data. It is critical to avoid duplication of the data in core business entities, especially when integrating with multiple source systems of similar nature (e.g. multiple ERP or CRM systems).

Start identifying the master data entities relevant to business cases. Define roles and responsibilities for business and engineers in managing the data entities (e.g. global MDM lead, local MDM owner, local data custodian).

Finalize business workflow for managing master data and reference data. Get the business team to define a master data template for core master entities and relevant mapping rules.

Ingest master data entities from all possible active source systems and centralize all data into the curated layer. Build the pipelines to prepare and maintain the master data with a global identifier. The global identifier is then further used and appended to all transactions data coming from corresponding source systems. Enable traceability of the master data to the source system from where the master data was ingested.

Step 6: Ensure that Data Governance is a Key Driver

A trusted SSOT platform requires embedded data governance. Ensure the selected solution addresses auditing and traceability, metadata management, data classification, data catalog, data lineage, data discovery, information lifecycle management, security and DevSecOps and consent management (for GDPR and other compliance requirements).

Taking an engineering approach covers all the dependencies for core data quality which enable end-to-end automated testing and higher productivity. This self-service engineering environment creates consistency, speed, and scale.

These capabilities are critical for enabling self-service and automation for data stewardship and data access. Additional scenarios that should also be considered depend on the industry and security/compliance requirements. Snowflake, with its out of the box capabilities like data encryption and dynamic data masking, is leveraged to address some of these compliance requirements.

Step 7: Enable Self-Service for Business Domains

Successful SSOT has business taking ownership of creating and maintaining data products with the support of the technical team. To do so, users need self-service capabilities for data governance (data classification, data catalog, data discovery and DevSecOps).

To enable efficient self-service, define role-based access control. Map access control according to data classification to ensure access to approved data sets. Define a policy and workflow for the data stewardship and automate the same through DevSecOps/IAAC. Install data discovery tools to help users find the data sets available for consumption and use the automated data stewardship workflow to request the data.

Once the setup is complete, users can start leveraging data tools like Dataiku, Alteryx or similar tools for data exploration and AI for self-service data consumptions. They can also ingest and transform the data using a no-code data pipeline solution like Calibo.

> "The self-service capabilities for engineering in terms of both the core data and analytics at the domain level assist users."

The self-service capabilities for engineering in terms of both the core data and analytics at the domain level assists users in creating and managing their data products obtained via the trusted data consumption layer. The data products created can be accessed by others through the defined data governance and DevSecOps principles.

Most business users rely on a consumption layer that is easy to understand and to work with the data. Enabling quick, easy availability of trusted data for consumption by various personas within the organization accelerates SSOT adoption. Other personas, apart from the data consumption layer, would also expect different services to be enabled for different use cases.

Step 8: Build a Business-Friendly Data Consumption Layer

Strategize on bifurcations for different business domains in line with security requirements. Get common data sets (e.g., customer data, vendor data, product information etc.) available from a common area.

Create the tables for business consumption using dimensional modelling and appropriate data pipelines to populate these tables. Create tables to store the output of analytics job or AI/ML models to enable easy access to other business users and enable business functions to take ownership of tables for specific business domains (based on the principles of Data Mesh powered on Snowflake).

Step 9: Build a Data Publishing Layer

Another important aspect of the platform is to create a capability for periodically publishing data to downstream applications or external actors.

Analyse various downstream application's dependency on the data from the platform and the desired interface to interact with the data. Typical patterns are text-based export (CSV/TSV), FTP-based data egress, REST API-based interface, Pub/sub-based data integration, and shared data sets (enabled by Snowflake's secured data sharing).

Prioritize the different patterns and build an egress capability accordingly, either as a periodic batch job or enabling reverse ETL tools like Hightouch, Rudderstack etc. Enable a REST API and pub/sub model for making the data available for downstream systems.

Step 10: Enable Sharing of the Data Assets

Create roles according to different users or types of users. Enable dynamic data masking on the datasets to be shared, based on the roles identified. Create data shares through Snowflake's secure data sharing for both internal and external users. Additionally, for external collaborators, enable Snowflake's data clean room for seamless collaboration.

In summary, creating an SSOT platform can be a daunting task. Taking a bite-size approach to building the capabilities incrementally helps to build users' confidence in the platform. Engage business teams early and continuously to build the platform. Coach and empower them to take ownership of the areas focusing on business domains and continuously augment the SSOT platform with data products/assets.

Case Study: Accelerating Data Products and Platform Adoption

A leading pharma company was on a journey to data platform modernization in the cloud. However, there was an overwhelming dependency on the technical team to cater to multiple uses cases, requested by different business functions.

A digital data/product engineering methodology helped the organization to accelerate the implementation while significantly improving the adoption of the platform by enabling business functions to build and own their data products with limited dependency on the technical team. A goal at the enterprise level was to create consistency, speed, and scale.

Execution Steps

Step 1: Identification and Finalization of the Source Systems

The technology team worked with the business team to identify a list of potential use cases with desired business outcomes. The use cases were analysed to identify the potential sources for the data, which were further analysed to find patterns.

Step 2: Finalizing the Designated Data Layers

The platform build was already in progress. However, a strategy for data layers was not defined and aligned. The team decided to go with data mesh principles on Snowflake. Existing objects were aligned to the new data layers.

Step 3: Refactoring Data Ingestion as Reusable Component

The existing data ingestion was a point-to-point integration. The technical team refactored the ingestion into a framework with reusable components as per the identified patterns. This led to a self-service engineering environment enables consistency, speed, and scale.

A common logging framework and alerting mechanism, along with exception handling, was added to the framework. Existing code was converted into reusable a component.

Step 4: Build Business Product for Early Review of Data

The product prototype was created in collaboration with the business team, which helped the technical team to iterate through the data refinement process until the veracity of the data was certified.

Step 5: Defining the Data Quality Framework

The core technical team then used the feedback and the learnings from the refinement process to create the data quality framework. Data standardization routines were developed as a reusable component for different data domains. Domain specific business rules were incorporated for different data layers.

Step 6: Managing Master Data as a Core Capability

For the existing master data entities already available in the platform, a master data governance model was defined by identifying key stakeholders for different subject areas. Master data owners defined the harmonizing and mapping rules for each master data entity.

The technical team worked with the global and local MDM lead and owner to implement the harmonization rule and created the master data repository to ensure that there is no duplication.

Step 7: Data Governance Framework Delivered

The technical team and business team worked together to classify the data in the platform and created the role-based access controls according to the data classification. Subsequently, the technical team enabled a metadata management layer as part of the core. Tibco EBX was used to create the data catalog.

The technical team leveraged Spline to capture the data lineage. A knowledge graph was created using Neptune. Custom data discovery was enabled on Neptune to provide technical and business domain specific details to both engineers and business users.

Step 8: Enable Self-Service for Business Domains

Business teams were enabled to bring their own data sets into the platform, enabling business ownership for domain data. Additional tools like DataIku were also integrated with the data platform to provide data munging capabilities. Tools to enable exploratory data analytics were also integrated.

Using DevSecOps enabled access to data sets (after data discovery) as automated data stewardship was constructed. This leveraged information from data classification and the role-based access control to ensure the user requesting a particular data set was authorized to consume the same.

Step 9: Business-Friendly Data Consumption Layer

Conformed dimensions were created as part of the core layer. The technical team (data architects and data modelers) enabled business domain specific data models using dimensional models.

Business teams, with support from the data engineering team, built the domain specific data pipelines using reusable components to populate the domain specific consumption layer tables. Data products were created by business teams for specific metrics and KPIs.

Step 10: Building a Data Publishing Layer

Reporting and BI solutions were integrated to the domain-specific models using appropriate connectors. Additional functionality for periodic extract for domain specific data sets were provisioned for downstream applications.

Step 11: Enable Sharing of the Data Assets

Snowflake's secure data sharing was leveraged for making the data products/assets available for internal users from other business functions.

Case Study: Accelerating Financial Reporting Outcomes for a Global Automotive Supplier

A leading global automotive supplier has operations in more than 50 countries and across 150 legal entities. It faced elongated book close and consolidation cycle time of about four work weeks, and sub-optimal use of finance teams due to the manual nature of collating data and generating financial information. In the previous chapter the business specific engagement drivers were highlighted for a global automotive supplier. This chapter focuses on the engineering methodology used.

Execution Steps

Step 1: Finalization of the Source Systems

With a focus on financial reporting, the pilot was identified for a specific region and four SAP systems were earmarked for the MVP implementation. While the team initiated the build, the business team worked on outlining the plan for the next phases in terms of regions and the applicable source systems. Overall, more than 10 ERP systems were identified for the final solution.

Step 2: Defining the Data Layers

With the principle of the incremental build for the platform, the team initially outlined three data layers: a landing layer, master data layer and consumption layer. In line with Snowflake and data mesh principles, the consumption layer was focused on the domain specific data sets/data products for finance.

Step 3 and 4: Implementing Data Ingestion and Data Product for Review

Reusable data ingestion pipelines were created for financial tables from an SAP system. Similar pipelines were created for master data entities from SAP. The business team analysed the data quality and provided feedback for data cleansing.

Step 5: Enabling Data Quality Measures

Guidelines provided by business were used to validate the data quality for different layers. Quality checks were integrated in the ingestion pipelines.

Step 6: Managing Master Data

The master data governance model was defined by identifying key stakeholders for different roles. Workflow was defined for managing the master data at the source layer. Data quality checks and rules were defined for the master data.

A global template was created for the ideal chart of accounts. Business established the harmonization and mapping rules for master data. The same was leveraged by technical team to populate the master data layer with golden records for identified master data entities, including the chart of accounts (mapping COA from each local unit to the global COA template).

Step 7: Implementing Data Governance with Ownership from Business

Metadata management and data catalog were implemented. The business team provided the guidance while the technical team supported the implementation.

Data lineage and process lineage were documented and integrated. Data discovery was enabled over the data catalog and metadata.

Step 8: Enabling Self-Service for Business

The business team was enabled to bring in additional data sets on a need basis. Guidelines were implemented to ensure high quality of the self-ingested data, leveraging the data quality framework. Tools were implemented to enable EDA for the business.

Step 9: Seamless Data Consumption

Business and data modelers collaborated to create a unified financial model for standardizing the financial data. The business team, supported by the technical team, built data transformation routines for populating the unified model.

Secured access controls were implemented on the data products based on the user's roles and departments.

Steps 10 and 11: Data Publish Layer and Data Sharing

As the financial data was classified as restricted, the data publish layer and the data sharing options were restricted to certain master data entities.

Support for API-based access and data extracts were enabled on the identified master data. The COA was not made available for general access.

A roadmap for additional capability was created for other patterns in line with the use case backlog. The driver was to incrementally augment capabilities according to the use cases.

Outcomes Achieved

- 50% reduction in book close cycle time.
- 30% improvement in resource productivity.
- 80% automation in consolidation of financial transactions across the enterprise.
- 100% data traceability and auditability.

Case Study: Improving Platform Trust and Adoption for a Leading Media Company

A leading media company has invested heavily in a modern data strategy leveraging the cloud for data, AI/ML and analytics. The data strategy was revised to use a hybrid cloud architecture with a focus on using Snowflake as the data core.

However, the adoption of the platform among business users was low. The following methodology was leveraged for improving the adoption and the overall platform capabilities.

Execution Steps

Step 1: Finalization of the Source Systems

The existing platform already had integrations implemented for data sources that are common to many business functions. External data feeds were also integrated.

For the Snowflake adoption, business and technology teams collaborated to create a phased roadmap for data migration and refactoring data pipelines for the existing data sets.

Step 2: Defining the Data Layers

One of the key aspects of the revised data strategy was to incorporate a data mesh on Snowflake and thus, the data layers on Snowflake were created to leverage the Snowflake best practices.

A core data zone was created to store the common data sets, while domain-specific data layers were created for business functions.

Steps 3 and 4: Implementing Data Ingestion and Data Product for Review

The existing platform had already implemented a reusable data ingestion framework. This was refactored to accommodate Snowflake as a new target.

Data from the existing platform was migrated to Snowflake in a phased approach. The business team supported the data validations and provided guidelines for integration with additional data sources into Snowflake.

Existing AI/ML models were deployed as engines for reusability.

Step 5: Enabling Data Quality Measures

An existing data quality framework was modified to work on Snowflake. New quality measures were added for additional new data sources.

Step 6: Managing Master Data

The existing platform had limited master data governance and implementation capability. While migrating to the Snowflake platform, the MDM capability was introduced in the new data strategy.

The technical team and the business team collaborated to implement the MDM component.

Step 7: Implementing Data Governance with Ownership from Business

The exiting stack lacked comprehensive data governance, which led to minimal adoption of the platform among business users as it struggled to establish the trust in the data to drive ownership.

In the new data strategy, additional functions were created, supported by business. These included design authority, architecture review board, demand management and prioritization, configuration and release management, and product management.

The Snowflake-based platform was fortified with additional capability in the form of metadata management, a data catalog and data lineage along with data discovery.

Step 8: Enabling Self-Service for Business

Tools were implemented to enable EDA for business. Additional tools were integrated for data munging.

Business teams were empowered to integrate domain-specific data into their respective zones.

Step 9: Seamless Data Consumption

A domain-specific data model wase created, using dimensional models, for business consumption. Core master data and conformed dimensions are managed and owned by the core data team and leveraged by business domains as read-only access.

Steps 10 and 11: Data Publish Layer and Data Sharing

Multiple egress patterns were implemented for various use cases.

Data sharing was enabled between core data layer (for conformed dimensions and master data) and the domain-specific layers for seamless data usage.

Chapter 3

Smart Analytics

Chapter Highlights

- Create a self-service environment for data and analytics at domain level.
- Establish data as a core asset.
- Empower business users to build and consume a domain level mega-dashboard.
- Use analytics, AI/ML for further growth through predictions and core asset enhancements.

In the previous chapter, we looked at why SSOT is critical, and the various steps to take to build out the SSOT. In this chapter, we will review some of those concepts and then look at the steps needed for generating insights from the SSOT.

"Data quality, master data etc. are critical for generating reliable insights."

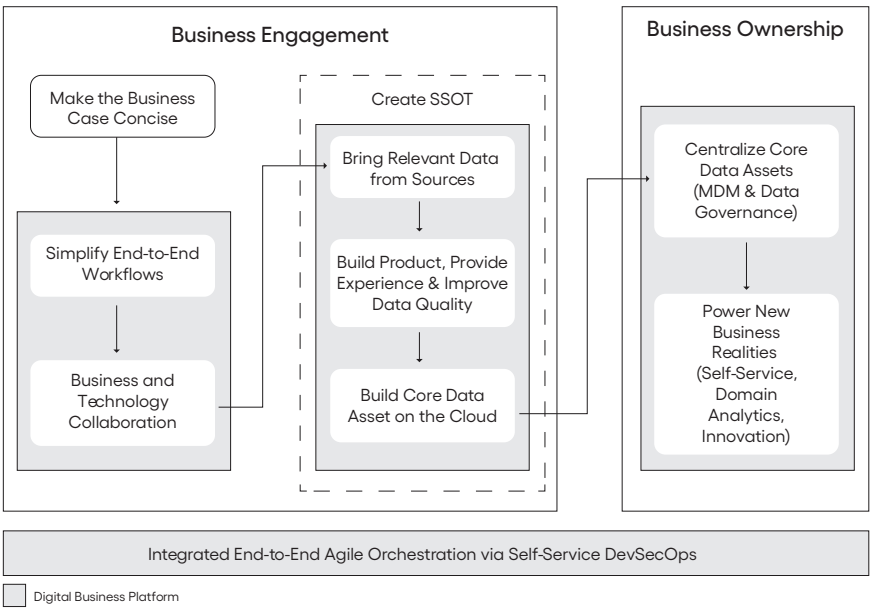
Data quality, master data etc. are critical for generating reliable insights. Challenges like missing data, master data not aligning across regions/products etc., can lead to incorrect insights. Data governance is therefore a critical stage in the overall Digital Business Methodology.

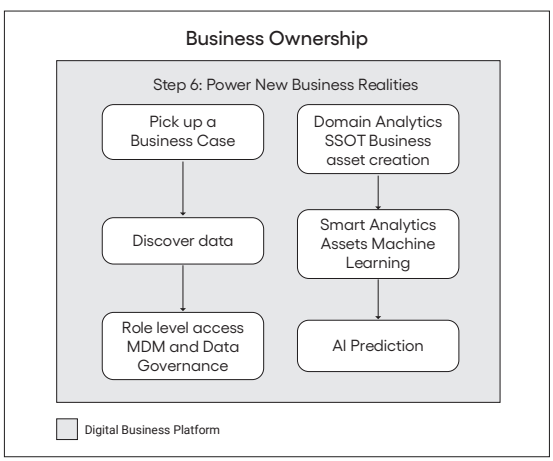
Many organizations are at some stage in their digital business journey. This journey needs to be governed by some basic principles, outlined below.

CHAPTER

SSOT and Culture

- Business outcomes with an incremental bite-sized approach.
- Business engagement/ownership Agile culture (SSOT–core data assets).
- Business intuition—SSOT at outcome level.





Regardless of where your organization is, if you are a data scientist, an analyst or a business stakeholder and would like to have some quick wins without sacrificing the long-term goals, here are the steps.

Business Methodology

Step 1: Pick a Business Use Case

Identify and prioritize a problem that you need addressed as soon as possible. Make sure it is not a predictive/prescriptive analytics use case.

Step 2: Discover Data

Get relevant data ready to address that use case. Understand the data from a business perspective. What does each attribute stand for? How is it interpreted in a business setting?

Step 3: Prepare Data with Master Data and Data Governance

You might need to clean up some master tables. Do that. Do not go overboard and launch into large multi-year MDM programs by derailing the business needs. You will get there incrementally. Implement data governance activities like data quality, role/group level access, encryption of sensitive data, etc. Now you have a small SSOT data set for this use case. This, by itself, is an asset owned by the business.

Step 4: Create Mega-Dashboards with Self-Service BI

Build some simple descriptive business mega-dashboards, an SSOT for the enterprise, on this data. The tools used could be whatever your firm has already invested in, such as Tableau, PowerBI, Qlik etc. In this stage, it is critical that you look for tangible business value related to the business use case that you started out with. You might use some AI/ML too, but you can still get by without sophisticated algorithms.

This is also an important stage to bring in any relevant external/market data, to augment the SSOT and analytics.

Step 5: Smart Analytics

If the above steps are done well, you would have generated some business value. This is what we call "bite-sized business outcomes." Move forward and leverage ML and statistics to dig deeper into the diagnostics of the data, and start establishing causal relationships. These insights can be generated in batch mode as well as real time (streaming analytics) mode.

Step 6: AI Prediction

Once you have established business value with such use cases, move on to more sophisticated ones which address the future predictive and

prescriptive analytics, using AI based algorithms.

> "The value of raw data becomes multi-fold when mega-dashboards, self-service BI, ML and AI are used to mine more insights from that data, thus impacting the revenues and margins."

AI can be used to enhance the core data asset. The value of raw data becomes multi-fold when mega-dashboards, self-service BI, ML and AI are used to mine more insights from that data, thus impacting the revenues and margins.

These steps will be repeated for multiple use cases, so they are iterative in nature.

As you start building out these business use cases, make sure there is a parallel MDM and data quality program running that organically keeps building up the SSOT.

The Result

The above approach of smart analytics ensures that you are addressing the need for engineering discipline in building an SSOT, but without sacrificing the "here and now" business imperatives.

This presents a bird's eye view of the methodology of smart analytics and can be quickly implemented using low-code/no code platforms like Calibo.

Case Study: Global Ingredients Manufacturer Increases Revenue, Enhances Cross-Sell and Upsell Recommendations

Overview of the company

A global ingredients manufacturer providing speciality chemicals for flavors, etc. to some of the world's largest companies in the food, perfumes, pharma and related sectors.

Targeted business outcomes

Increase revenue for top 25 accounts:

- Review synergy and product association (complementary products).
- Next product to buy algorithms.

- Bundling/product mix recommendations.
- Increase order value and grow revenue by customer across all categories.

Cross-sell recommendations (increased order value and basket size):

- Recommendation of products and ingredients to cross-sell a wider portfolio (within and outside business units) to key accounts to increase basket value.
- Maximize value to customers and increase revenue without the recurring cost of marketing/sales.

Upsell recommendations (increased order value and basket size)

- Recommendation of products and ingredients to upsell (higher value and quality products) to key accounts.
- Filter by region, business units, sub-business units, key accounts.

Execution Steps

Step 1: Understanding the Business/Problem

- Document the current cross-sell/up-sell process, and the key challenges.
- Set up the data sandbox with entity relationships, and, flattened /de-normalized data sets, as needed.
- Outlier analysis and inclusion/exclusion.
- Finalize the kick-off feature-set, post feature selection.

Step 2: Data Discovery

- Review the data availability, transactional as well as master data.
- Understand the entity relationships and attributes, as related to cross-sell/up-sell.
- Evaluate the level of data quality, for analysis.
- Identify missing values in data and plug the gaps.
- Initial round of data cleansing.

Step 3: Role-Level Access/MDM/Data Governance

- Review the current product master and build a global product master to get a single view of product across the company with clear alignment of attributes and definitions, to drive synergies across products and ingredients.
- Develop SSOT.
- Potential for leveraging data from SSOT to quickly solve business challenges.

- Opportunity to establish higher veracity in data and improve on current limited data governance.
- Identification of reporting consumption patterns.
- Further refinement of data quality checks.
- Data stewardship and governance strategy.
- Tools and frameworks for data quality, glossary, data lineage etc.
- Data discovery, accessibility, security.
- Deployment and provisioning strategy.

Step 4: Analytics SSOT, Mega-dashboards, Business Ownership, Asset Creation

Analyse opportunities for growth.

- Study the data using histograms, bar-charts, heat-maps etc.
- Augment using relevant external/market data.
- Build mega-dashboards to bring the insights together.
- Check for correlations between the features.
- Finalize the missing value imputation logic.
- Convert categorical/text values to numeric.
- Standardize/scale the features.

Step 5: Smart Analytics Assets, ML, Patterns and Behaviours with Real-Time Data

- Build ML models based on the inputs from the previous steps.
- Iterate to the data review step, if required.
- Pick the ML model that best suits this particular data set and context.
- Generate preliminary results.
- Technical evaluation of the model/results, testing.

Step 6: AI Prediction, Appropriate Algorithms for Predictions, SSOT/AI as Asset

- Use AI for predicting market trends for upsell/cross sell opportunities.
- Augment the above models with market information from additional data sources.
- Build comprehensive AI models for predictive analytics.

Outcomes Achieved

Using step 4:

- Eight established patterns.
- 70-plus recommendations.
- About $30 million potential opportunity.

Using steps 5 and 6:

- About 700 recommendations.
- About $300 million potential opportunity.
- Case Study: Healthcare Company Shifts Data and Legacy App to Cloud and Sees Improvement in Data Availability and Reduction in Data Maintenance Cost
- Overview of the Company
- A global top five healthcare company that focuses on discovery, development, manufacturing, distribution, and commercialization of pharmaceutical products.

Targeted Business Outcome

Migrate data and legacy application from on premise to cloud.

- This includes data related to trials, supply chain, testing, patient records, cost of treatment, human resources, suppliers, marketing, finance, etc.
- Currently, Phase III clinical trials alone generate an average of 3.6 million data points, or three times the data collected by late-stage trials 10 years ago.
- Data is the cornerstone of the pharma industry. It ensures the efficacy, quality, and safety of products, supports R&D, drives innovation, and helps meet regulatory requirements.

Reduce the cost of maintaining the data.

- The pay-per-use model of cloud consumption leveraged to deliver significant cost reduction in data management.
- Making data sharable to an increasingly decentralized workforce without investments in additional data centres in remote locations.

Future-ready with the ability to leverage AI and ML

- Enhance security through fine grain access across various layers.
- Apply AI and ML for further business solutions like foot optimiser and patient enrolment solutions.

Execution Steps

Step 1: Understanding the Business/Problem

The client from the pharma industry had stored about 10 years of its data in multiple on-premises environments with an Oracle database. The data was stored based on past industry best practices. With rapid evolution in the industry and the needs of the organization changing, the client wanted to migrate its Clinical Analytical Function Database (CAF DB) to the cloud.

This is data from multiple sources related to global drug discovery, development, and trials across various sources are collated and unified in unification layer of the platform. The client also wanted six Qlik based applications migrated to the cloud along with the data.

Step 2: Cloud Transformation

AWS was the cloud platform to which the data and applications had to be migrated. The client named its multi-layered cloud system as the data platform. The data was moved by the team from the pre-landing layer to the landing layer and then the unification layer before being handed over to the publishing layer in the Snowflake database for further presentation and any analytical needs.

Step 3: Build Extract, Transform and Load (ETL) Pipeline

To move data from an on-premises database to Snowflake, an ETL pipeline was built. There was adaptation required for the existing application architecture and code base to be completely rewritten, to ensure that they are functional using the new data platform and the principles on the new platform. Data layers were created to hold numerous types of data needed to compute KPIs.

Step 4: Utilization of Snowflake

Altimetrik helped in unifying the data by using a unified model agreed across the organization in the unification layer and further enhanced the data using the refinement and publish layers with the help of snowflake. Snowflake's features of fine grain access control helped in providing access across the data layers and the various personas.

Step 5: Accelerated DevOps

Effective collaboration with infrastructure and the DevOps team for various pilots

and Agile rollouts. Setup iterative feedback loops along with the various teams and functional groups. Automated metrics for productivity and quality.

Step 6: Qlik Sense for Data Visualization

Altimetrik then created a semantic layer on the data with Qlik Sense, from where the client's team could create visualizations for the senior management. In addition to the Qlik Sense integration, Altimetrik broke down the code for all six client applications written in Oracle MySQL, transformed, and migrated them to the new environment.

Outcomes Achieved

- Two times improvement in data availability.
- Reduction in data maintenance cost.
- Robust collaboration and operational excellence.

Case Study: Food Company Gains Predictive Insights on Consumer Buying Patterns and Revenue

Overview of the Company

A leading food brand in Australia had its bakery operations spread across multiple countries in the Asia Pacific region. It sells biscuits and cookies through mass merchandisers and retailers.

Targeted Business Outcomes

- Provide predictive insights on consumer buying patterns and consumer behavior.
- Establish understanding on competitor strategies.
- Develop business insights on revenue, profitability, and cost.

Execution Steps

- Discover consumer insights, business insights and data platform.
- Finalize metrics and design.
- Finalize data architecture.
- Align panel and scan data periods.
- Data modeling, master and reference data, data governance, data security and compliance.

- New pipeline for Snowflake.
- Finalize business insights mega-dashboard for self-service business intelligence.
- Apply ML and later AI for repeat purchase, substitutability, purchase behavior. modeling, flavor/trends analysis, correlation analysis and competitor insights.

Outcomes Achieved

- The company built a successful consumer clustering algorithm to gain a richer understanding of consumers and buying behavior; seven new clusters were identified.
- Consumer clustering in turn informed consumer flavor and product substitution preferences.
- It was able to successfully corelate competitor pricing strategies on product sales.
- The firm developed detailed business insights on revenue, profitability, and cost at a category level.

Case Study: Agro Company Deploys Yield Forecasting Solution and Significantly Improves Accuracy

Overview of the Company

A global Agro company providing farm produce to some of the world's largest retailers selected Altimetrik for implementing a yield forecasting solution to provide accurate yield forecast across its harvesting cycle for their green houses.

The solution included deep analysis of historical data, review of the data patterns, and developing a forecasting engine with historical time series data and regressors, to deliver the yield forecasting results. In addition to the forecasted value, the solution also consisted of user interface (UI) dashboard for business to review and analyse the yield patterns across various parameters to help plant managers and growers make the right decisions.

Targeted Business Outcome

- Optimal demand management by improving yield forecast accuracy to 90%.

Execution Steps

1. Requirement analysis for UI/dashboard with business.
2. Acquire the historical data and data analysis.
3. Automated data pipeline development.
4. Data accuracy testing.
5. Identification of data patterns to develop an ML model.
6. Forecast model improvement leveraging growers' experience.
7. Development of user interface dashboard for business users.
8. Unit testing and user acceptance testing.
9. Postproduction environment construction.
10. Deployment of the solution.

Outcomes Achieved

- One-week yield forecasting accuracy went up to 92%.
- Six-week yield forecasting accuracy went up to 86%.

Chapter 4

Digital Engineering Ecosystem in a Customer Environment

Chapter Highlights

- End-to-end agile, digital, self-service ecosystem with tools optimization, to enable quick adoption to arrive at SSOT.
- Cloud agnostic optimization that provides performance optimization, risk optimization, cost optimization, and future proofing.
- End-to-end data governance for managing and ensuring the quality, integrity, and consistency of data across an organization.
- Shift Left principles with continuous quality and security engineering adoption.
- Continuous monitoring that adopts principles of site reliability engineering to provide uptime through high availability.

While we focused on the creation of SSOT with self-service domain analytics with AI in the previous chapter, in this chapter we examine the importance of looking at the digital engineering ecosystem that supplements and complements Snowflake to accelerate business outcomes.

As many organizations reinvent their business models with a focus on innovation and customer centricity, it's crucial for them to develop flexibility and responsiveness. Digital businesses need to stay ahead of the curve by continuously learning and adapting to changing market conditions.

Figure 1 below outlines the flow of the digital business methodology with a focus on the underlying digital ecosystem, which is based on the DevSecOps principles that act as the fabric to connect systems needed to build the isolated SSOT environments.

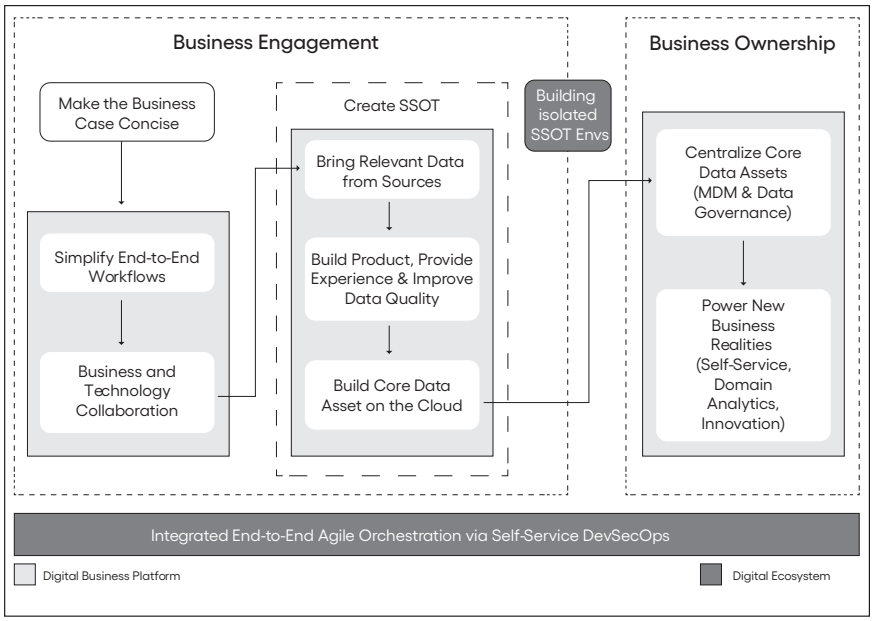


Figure 1: Digital Ecosystem for adopting Digital Business Methodology

> "As many organizations reinvent their business models with a focus on customer centricity, it's crucial for them to use automation and self-service capabilities to develop flexibility and responsiveness"

The digital business methodology focuses on business agility via a self-service culture structured and enforced by foundational components across the SDLC that must be developed or enhanced to achieve a frictionless engineer experience. This allows organizations to focus on building innovative solutions rather than overloading time and resources on managing technical debt.

The diagram below illustrates the step-by-step approach to enable the digital ecosystem.

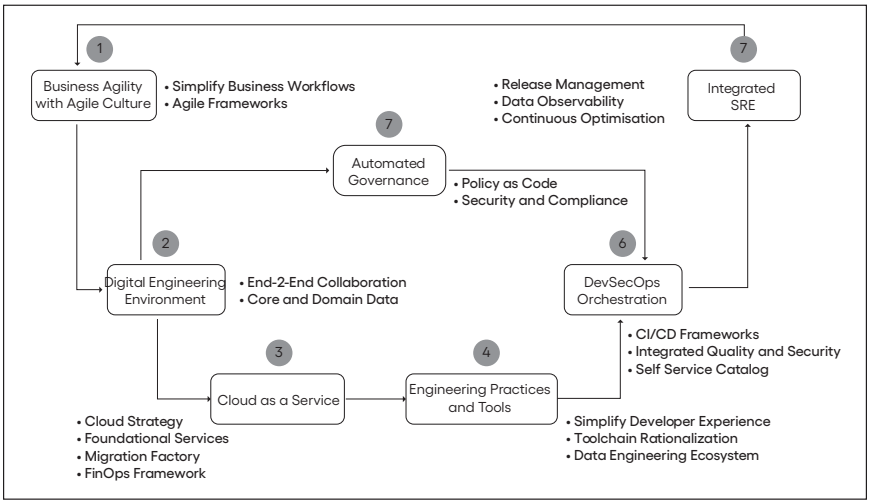


Figure 2: Step-by-step approach to enable the Digital Ecosystem

Simplify the Business and Get Started with Agile Culture

To succeed in today's fast-paced business environment, companies must be proactive in meeting customer demands and speed to market. As digital technologies become ubiquitous and essential for many roles, Agile fluency is becoming increasingly important to build an effective digital business.

Developing Agile fluency allows teams to improve their productivity, collaboration, and communication skills, ultimately enhancing the overall effectiveness of the digital ecosystem. With a focus on the overall business process instead of siloed data, teams can refine processes, achieve better quality data, and be more responsive to user needs for analytics.

Step 1: Simplify Business Operating Models and Workflows

Evaluate the organization's readiness for Agile adoption by assessing the existing culture, processes, and mindset. Identify any potential challenges or resistance to change and develop strategies to address them. Cultivate a culture of collaboration, flexibility, and continuous improvement to support Agile practices.

Form cross-functional teams composed of individuals with diverse skills and expertise relevant to the project or initiative. These teams should be empowered and self-organizing, with clear roles and responsibilities defined.

Establish a process for prioritizing work and managing product backlogs

effectively. Define criteria for prioritization, involve stakeholders in the process, and regularly review and reprioritize strategic initiatives and product backlog items based on changing business needs and customer feedback.

Step 2: Adopt Agile Frameworks

Select and implement an Agile methodology that best suits the organization's needs, such as Scrum, Kanban, or SAFe. Provide training and support to ensure that team members understand and embrace Agile principles, practices, and ceremonies. Adapt the chosen methodology to fit the organization's unique context and requirements.

Embrace an iterative and incremental approach to software development or project execution. Break work into smaller, manageable increments or sprints, with defined goals and deliverables. Use techniques such as user stories, epics, and sprint planning to manage and track work progress. Conduct regular retrospectives to reflect on lessons learned, identify areas for improvement, and adjust to optimize future iterations.

Implement continuous integration and delivery to streamline development and validation activities. Automate build, testing, and deployment processes, enabling frequent and reliable releases. Establish metrics and key performance indicators (KPIs) to measure progress, track team performance, and identify areas for improvement.

Digital Engineering Environment: End-to-End Collaboration

To successfully implement Agile frameworks for data platform development, teams must take proactive steps to ensure that they are able to deliver high-quality data products and insights in a collaborative manner. By doing so, they can improve their operational efficiency, increase customer satisfaction, and gain a competitive advantage in the marketplace.

Step 1: End-to-End Collaboration

Foster a culture of open communication and information sharing across teams and departments. Encourage regular brainstorming sessions and knowledge-sharing forums. Provide channels for effective communication such as shared document repositories, wikis, or slack channels.

Implement digital tools and platforms that facilitate communication, collaboration, and transparency. This might include product management tools, Agile boards, chat platforms, virtual collaboration tools, and version control systems. Ensure that these tools are easily accessible, user-friendly, and integrated to promote seamless collaboration.

Step 2: Data Democratization with Core/Domain Data

Data must be accessible and available to a broad audience within an organization, empowering individuals to access, understand, and use data for decision making and problem solving.

Identify the core data and domain-specific data that are essential for your organization's data engineering initiatives. Define the data models that represent the structure, relationships, and attributes of the data. Determine the appropriate data storage and infrastructure required to support the core and domain data. Provide mechanisms for data engineers, data analysts, and other users to access and analyze the core and domain data.

Develop a data governance framework that defines policies, procedures, and guidelines for data access, security, quality, and privacy. Identify relevant data sources within the organization and create a comprehensive data catalog that provides information about available datasets, their definitions, and usage guidelines. Set up processes and tools for integrating and preparing data from different sources. This involves data cleansing, transformation, and aggregation to ensure data accuracy, consistency, and usability.

Cloud as a Service

Organizations seeking to enhance delivery speed and promote team agility are increasingly embracing multi-cloud initiatives to cater to the ever-evolving data platform and infrastructure demands. Cloud as a Service (CaaS) offers organizations numerous advantages, including scalability, cost efficiency, rapid deployment, focus on core competencies, global accessibility, enhanced security, and opportunities for innovation.

Step 1: Cloud Strategy

Define a cloud strategy for data platforms that ensure data initiatives and investments are aligned with overall business objectives. Outline clear goals and objectives for data management, analytics, and insights generation, ensuring that the chosen cloud data platform supports these objectives effectively.

Familiarize yourself with the principles and objectives of the Cloud Smart strategy. Recognize that it goes beyond a simple "lift-and-shift" approach transitioning to self-service cloud templates. The cloud strategy should outline guidelines for robust access management, security measures to ensure relevant compliance with relevant regulations and standards, and data integration standards and data quality validation to create a unified view of data across the organization.

Step 2: Foundational Services

Outline cloud foundational services that provide the necessary capabilities,

tools, and infrastructure to discover, store, process, integrate, and analyze data, enabling organizations to derive valuable insights, make data-driven decisions, and drive innovation. There should be a structured process to qualify cloud services, organized workspaces for security and cloud workload management, and adopted service control policies to enforce security controls.

Establish processes and mechanisms to control and manage user access to data and system resources within a data platform. Automate the process of granting access to cloud services and systems. Define roles, permissions, and time-limited access to ensure that only authorized individuals can access resources.

Key aspects of entitlement management should support authorization and permissions, role-based access controls, data segmentation, access control policies, privileged access management, data masking, and access request / approval workflows.

Step 3: Migration Factory

Develop a cloud migration factory that follows best practices to ensure a smooth and efficient transition of digital assets to the cloud. The factory model should include workload discovery, target architecture blueprints, self-service migration frameworks, and continuous optimization. To minimize disruptions to business operations, take a phased approach, starting with a small pilot project before migrating large workloads and infrastructure.

Create a detailed plan for migrating data products and infrastructure to the cloud. This strategy should consider factors like data management, security, workforce training, and financial operations (FinOps). A well-designed cloud migration factory will help increase agility, enhance cost savings, and improve scalability by leveraging the benefits of cloud technology.

Step 4: FinOps Framework

A critical step is to develop a FinOps framework that focuses on achieving financial accountability and efficiency by allocating and attributing costs, optimizing expenses, forecasting budgets, and implementing transparency and accountability mechanisms.

Measure “golden signals” and other metrics to make more informed decisions about the cloud consumption model. Optimize FinOps practices to manage and optimize cloud costs. Monitor usage patterns, implement cost allocation, and identify opportunities for cost savings.

Additionally, analyzing cloud costs in the context of the business function supported is crucial to fully understand the value. By using data-driven spending decisions and tools like sensitivity analysis, you can maximize business value, gain better visibility into data platform usage patterns, and identify opportunities for cost savings.

Engineering Practices and Tools

When defining the optimal engineer experience, organizations need to take a holistic view into all the engineering environment and associated workflows. The end-to-end environment must incorporate best practices related to high-quality data sources access, automated setup of development environments, reusable libraries and frameworks, and deployment orchestration.

By integrating all these elements, teams can ensure that their data products and services are rapidly developed and delivered with high levels of quality and security. This end-to-end engineering environment is crucial for consistency, speed, and scale.

Step 1: Simplified Engineer Experience

Understand the specific needs and requirements of data engineers within your organization. Determine the tools and practices that are commonly used for data engineering tasks such as data ingestion, transformation, modeling, and integration. Redefine the data engineering experience by creating a suitable development environment that enables team members to work effectively and efficiently.

> "To succeed in today's fast-paced business environment, companies must have business engaged and an orchestration layer that enforced consistent methodology."

Conduct a process mapping exercise is critical to document the current engineering processes; pinpoint outdated practices, gates, and guardrails; and identify governance checks.

Create optimal engineering workflows, eliminating task-driven support activities and developing an integrated toolchain that provides full transparency, traceability, and continuous integration. Encourage the creation of a community to enhanced collaboration and knowledge sharing, standardization of best practices, faster time-to-market, and increased innovation.

As engineers are encouraged to contribute to shared codebases and collaborate with other teams and departments, silos within the organization can be broken down and communication and collaboration can be improved.

The primary objective of the engineer experience is to deliver a seamless and enjoyable experience which can enhance their job satisfaction, increase their productivity, and ultimately contribute to the success of digital business initiatives.

Step 2: Toolchain Rationalization

Review the existing tool chain landscape, adoption patterns and effectiveness of independent tool feature sets. Evaluate and select the most suitable tools for different stages of the data engineering process to ensure they meet the organization's specific needs and requirements. The toolchain framework should supplement and enforce developer workflows, ensuring transparency and traceability.

Build an enterprise toolchain that provides an integrated experience that includes end-to-end orchestration, starting with project onboarding, creation of engineer workspaces, automated builds and package management, and code quality and security analysis.

It is important to have readily available toolchain templates, which can be modified or updated by teams using a version control system. To facilitate this, incorporate pipeline stages, parameters, and gates as "pipeline as code." This allows for a more streamlined and automated workflow, as well as easier maintenance and consistency across the development process.

Step 3: Data Engineering Ecosystem

Based on the target state engineer experience and tool chain rationalization exercise, you will need to define the data studio that promotes best practices and guidelines to streamline data engineering workflows. And based on the identified needs, define the target tooling and technology stack for the data engineering studio.

Set up a centralized infrastructure environment for the data engineering studio that provides the necessary resources, scalability, and security features required for data engineering tasks. Implementation of self-service analytics tools and platforms allow users to explore and analyze data independently. The application of robust security measures protects sensitive data and ensures compliance of privacy regulations.

Automated Governance – Programmatic Policy Enforcement

Automated governance in a digital ecosystem is a method of using technology and streamlined processes to ensure that technology teams comply with regulatory and compliance requirements. It involves integrating policies and controls into the development and delivery pipelines to ensure that the business meets the necessary standards and regulations without compromising the speed of delivery.

By automating governance, businesses can reduce the risk of non-compliance and speed up the time to market, which can positively impact the bottom line.

Step 1: Policy as Code

A policy-as-code blueprint is a valuable tool for organizations to ensure that their policies and standards are consistently enforced across the product lifecycle. To

create effective blueprints, it is important to define policies in a clear and concise manner, use standardized formats, implement version control, test policies, and monitor them regularly. Involving all relevant stakeholders in the process can help ensure that the policies are effective and meet the needs of the organization.

Establish guidelines and guardrails related to delivery frameworks, engineering practices, data security and controls, compliance with regulatory requirements, production deployments, and systems operations. Develop a programmatic approach to defining and enforcing delivery, engineering, infrastructure, and compliance requirements as executables. Design policy blueprints and incorporate a policy engine to construct intelligent end-to-end data engineering pipelines.

Step 2: Security and Compliance

Develop security and compliance guidelines related to secure development practices, infrastructure and configuration security controls, security testing, access and identity management, security monitoring, compliance automation, auditability, and governance to minimize security risk and achieve compliance objectives while maintaining agility and rapid delivery of data products.

Use policy blueprints and automated governance techniques to enable teams to enforce their compliance with organizational best practices and guidelines in real-time. Implement automated governance mechanisms to reduce manual checkpoints and delayed response times. The automated governance model should incorporate approval workflows, security checks, and compliance monitoring within the self-service framework.

DevSecOps Orchestration via CI/CD

It is important to design end-to-end orchestration capabilities that support Agile delivery frameworks while also enforcing engineering code quality practices, test gates and guardrails, immutable infrastructure/environment provisioning and management, and integrated change management policies.

By integrating all these elements, organizations can ensure that their applications and services are rapidly developed and delivered with high levels of quality and security. With the enforcement of engineering code quality practices and test engineering gates, code can be thoroughly tested and meet the organization's standards for quality and security.

Step 1: CI/CD Workflows

The orchestration pipeline services as the foundation of an effective continuous integration (CI) and continuous delivery (CD) framework. Create end-to-end data pipeline orchestration that supports the entire lifecycle, including ingestion, transformation, testing, security, compliance validation, deployment, and observability of services.

Adopt automated deployment tools and frameworks to deploy the data engineering artifacts to the target environment. After successful validation, the data engineering artifacts—such as ETL scripts, data pipeline definitions, or configuration files—will be packaged for deployment.

Step 2: Integrated Quality and Security Practices

Integrate test automation capabilities related to unit testing, integration testing, performance testing and fault injection testing into the toolchain to reduce defects and increase system availability. Key aspects to be considered are continuous testing, environment management, test data management, quality gates, monitoring, and observability.

CI/CD practices should be used to integrate security testing into the pipeline, allowing for frequent and comprehensive testing. Implement static code analysis tools and techniques such as taint analysis and data flow analysis to identify potential vulnerabilities in source code. Leverage data encryption to protect data while it is in transit or at rest in the cloud.

Step 3: Self-Service Catalog

Understand the significance of self-service capabilities for both business and technology teams. This means providing them with the tools and frameworks they need to independently perform their tasks without relying on complex workflows, ticket-driven operating models, or ineffective support models.

Leverage DevOps building blocks and develop self-service frameworks to streamline and accelerate data product engineering, by providing an efficient way to access and manage needed resources. Develop a simple and user-friendly service catalogue that provides end-to-end orchestration and easy consumption. This catalogue should leverage reusable building blocks that eliminate waste and enable users to create solutions more efficiently.

Ensure that the self-service capabilities are effective by regularly monitoring and refining the platform based on user feedback and evolving business requirements.

Integrated Site Reliability Engineering

Giving engineers operational responsibilities has greatly enhanced the quality of the services they provide, both from a business agility and technology point of view. Site reliability engineering is an IT operations discipline that extends DevOps concepts by introducing an engineering approach to data products operations.

It introduces novel concepts such as programmatic capacity management, observability, toil detection and automation, monitoring and alerting capabilities,

pipeline health dashboards, and self-healing services to proactively address potential issues before they can negatively impact customer experience.

Step 1: Release and Incident Management

Effective release management practices for businesses require the CI/CD frameworks with release governance, using version control, automated release readiness checks, pre-deployment environment health monitoring, data pipeline promotions and continuous monitoring for issues in real-time.

Transition to an active service engagement model with integrated risk and compliance, release management, change management, incident management and problem management. To ensure the highest possible level of reliability, the CD framework should incorporate SRE gates, monitoring reliability measurements before and after a deployment event.

Step 2: Data Observability

Define SRE principles and practices related to resiliency, availability targets, scalability and capacity management, observability frameworks, and service level objectives. Develop standardized frameworks and templates for monitoring as a service, observability as a service, continuous delivery and deployment, and dashboarding.

Integrate reliability parameters with the use of various testing methodologies and tools such as load testing, performance testing, resiliency targets and trends, error budgets, and fault injection testing.

Integrate alerts with monitoring tools for effective incident management. Best practices for this integration include selecting a monitoring tool that supports alerting, defining clear and concise alert thresholds, establishing escalation policies, automating alert triage and resolution, integrating with incident management tools, and regularly reviewing and updating alerts.

Step 3: Continuous Optimization

Complete gap analysis of operational runbooks, and monitor and enhance automation framework and self-service capabilities related to self-healing, toil elimination, and fault tolerance remediation. Scale SRE adoption to other teams and services across the organization.

Continuously improve practices, tools, and systems based on experience and positive outcomes. Develop automated processes and services that can detect and resolve compliance or operational issues without manual intervention, reducing maintenance overhead and improving reliability. Use the feedback and data to drive innovation within the organization, improve team productivity, and enhance delivery predictability.

Case Study: How a Pharma Data Organization Enabled Business Teams with Enterprise DevSecOps

Overview of the Company

A leading multinational pharmaceutical corporation focused on innovative science and digital technologies to create transformational treatments in areas of great medical need.

Background

The company had embarked on multi-year journey to simplify and streamline the ingestion, transformation, and visualization of trail data to empower data scientists and accelerate innovation. The company made strategic decision to provide a modern platform including data engineering and DataOps pipelines as a service to its partner ecosystem.

Targeted Business Outcomes

- Improve results delivered to the business, including enhancing quality.
- Develop new capabilities and deliver solutions with increased speed and quality; increase release frequency by six times.
- Improve engineer productivity by 30%.
- Enhance IT resource usage, adopting a cloud infrastructure and engineering practices such as Agile and DevOps.

Approach

1. The approach focused on improving on transformational tenets such as:
2. Defining the business value of the transformation.
3. Preparing for culture change.
4. Starting small but being strategic.
5. Mapping out the technology platform approach including self-service.
6. Seeking out partners and expertise (four different divisions were considered).
7. Gathering feedback and making needed adjustments.
8. Scaling the transformation.

Tools Optimization

- Transitioned to open source and Atlassian toolchain.
- Standardized code analysis profiles with embedded gates.

- Created dynamic build farms for compilation of various tech stacks.

Integrated Workflows

- Re-engineered engineering data and DevOps workflows, gates, and guardrails to enable seamless CI/CD.
- Standardized on branching model, code review and analysis, and orchestration pipelines.

Qlik Sense Deploy

- Created custom methods for supporting multi-branch development, packaging, and deployment of solutions.

One Click Onboarding

- Platform to allow for self-service onboarding and entitlement management of new data related projects and data asset access management.

Outcomes Achieved

- 50% reduction in product creation from ideation to MVP.
- 30% engineering productivity.
- 15% reduction in Cloud consumption.

Case Study: How a Digital Media Company Embarked on the Digital Ecosystem Adoption Journey

Background

The company was seeking a migration of its broadcast media, digital and enterprise technology products from a private data center to a public cloud. As part of the transition, it was looking to improve its product engineering capabilities by standardizing the engineering environment, embedding quality and security throughout the delivery process, and incorporating continuous integration/continuous deployment frameworks while accommodating the nuances inherent in the different businesses.

Problem Statement

- Shadow Agile and DevOps were adopted with inconsistent tool chain design and adoption patterns, no standardized or enterprise level baselining on best practices and enforcement.
- Lack of visibility of engineering maturity, inadequate tool chain design patterns, and cross-team dependency management impacting quality and speed of scrum teams.

Targeted Business Outcomes

1. Standardize engineering environments, to avoid inconsistency.
2. Gain clear visibility of engineering maturity with cross-team dependency management.
3. Provide a predictability view of the ongoing development activities to bring in release time visibility.
4. Deploy automation and automated drift remediation to enhance engineering productivity with self-service engineering environment for enhanced productivity.
5. Deploy on-demand release framework with CI/CD gates/guardrails to support green releases with an audit trail, resulting in faster time to market.

Approach

1. Lead six teams in review and modernization of the product engineering practices environment strategy and tool chain.
2. Multi-track transformation consisting of:
    - Cultural environment for change adoption.
    - Engineering practices and workflows refinement.
    - Build/deploy automation.
    - Environment strategy to enable CI/CD/CT.
    - Environment right-sizing to help reduce infrastructure costs.
    - Tool chain rationalization to simplify the toolsets adopted.
3. Construct playbook for transformation.
4. Adopt playbooks, reusable artifacts, accelerators, and training of client teams—empowering them to self-organize toward achieving a higher state of maturity.

Outcomes Achieved

- Deliver productivity increased by 35%.
- Agile adoption reaches 80%.
- 5000 employees onboarded.

Case Study: Enabling 360° Consumer Insights for a Biscuit Manufacturer

Background

The company wanted to understand customer sentiments and perform a product analysis to arrive at accurate positioning and go-to-market strategies. It had to build a data foundation platform to consolidate, aggregate and analyze business information on consumer insights on both internal and external data. The company also need to extract business insights through advanced analytics to support business decisions.

Targeted Business Outcomes

1. A trusted data foundation with implementation of single source of truth.
2. Insights Cockpit with a unified view of overall activities with ease of searchability.
3. Creating isolated environments and scaling preproduction environment.
4. Seeding preproduction environments with production data.
5. Handling errors with the automated release of a database and related application code.

Approach

1. Creation of SSOT.
   Unify scattered sources of data to create a single clean source across sales, scan, customer, market.
2. Clear visibility across the value chain.
   Have the ability for experimental detection of business outcomes based on data patterns and interactions through ML principles.
3. ML Principles adoption.
   Adopting MLOps to train the model to bring in better prediction.
4. Improving release frequency.
   Fully agile delivery with extensive expert inputs for frequent delivery of tangible business value.

Outcomes Achieved

- Release frequency improved by 50%.
- Prediction improved by >45% .
- Near real time dashboards for immediate decisions.

Conclusion

To establish a highly efficient digital enterprise, it is critical to evaluate the strategies concerning organizational agility, cloud as a service, optimized engineer experience, DevSecOps frameworks, and integrated site reliability engineering. To enable a digital ecosystem, it's essential to overcome the traditional way of working that might impede progress.

A well-defined digital ecosystem can provide several business benefits, including increased efficiency, enhanced collaboration, improved customer experience, better data management, and faster time to market. By leveraging the digital methodology, organizations can achieve greater productivity, innovation, and competitiveness.

Chapter 5

Accelerating Snowflake Adoption via Calibo

Chapter Highlights

- Establish and integrate a digital business platform (cloud data engineering environment) to enforce Digital Business Methodology in a cloud agnostic ecosystem.
- Create a SSOT and data assets using Snowflake via an incremental bite-sized approach.
- Improve the experience of business users by creating, managing, and governing data assets through self-service analytics and data operations.
- Simplify the end-to-end workflows and enforce compliance and governance using automated guardrails.

The previous chapters explained the digital business methodology; how to get business engaged and take ownership in data activities such as SSOT, domain data and analytics assets and data governance. We also detailed how we create an ecosystem to simplify digital data and analytics product creation.

This chapter will describe how to apply the DBM and use Calibo as an end-to-end digital ecosystem to enforce it within the customer ecosystem and accelerate business outcomes.

Every enterprise is pushing to improve business results. Data is one of the key ingredients to new business models and insights in every industry. This is why Snowflake with its advanced cloud data capabilities is one the fastest growing technology companies in the world.

Often people associate data needs with storage and compute. However, data value is constructed through product creation life cycle and delivered in an

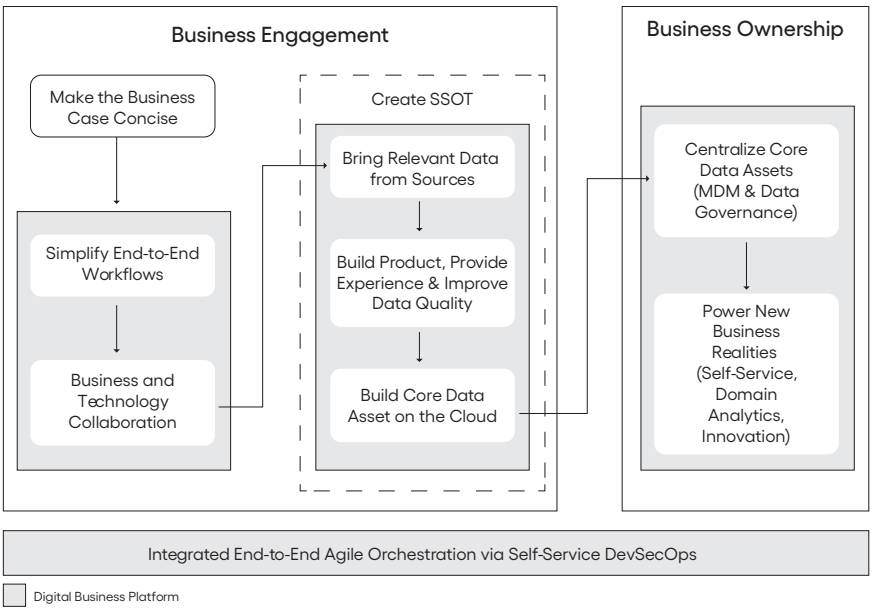
iterative manner.

Lifecycle methodology, DevSecOps, data operations and governance are key aspects organizations need to master when they try to get more value. Unfortunately, this is often neglected and left to technologists to support.

Business leaders and users and technologists need to work together and detail their path to success. They need to question themselves and address the following:

- Where is data sourced from and how it can be discovered and accessed?
- How is data enriched and how can it be reused?
- How can you keep data quality under control?
- Who can analyze data and how can they drive insights?

The below diagram explains digital product creation and how a digital platform such as Calibo simplifies it.

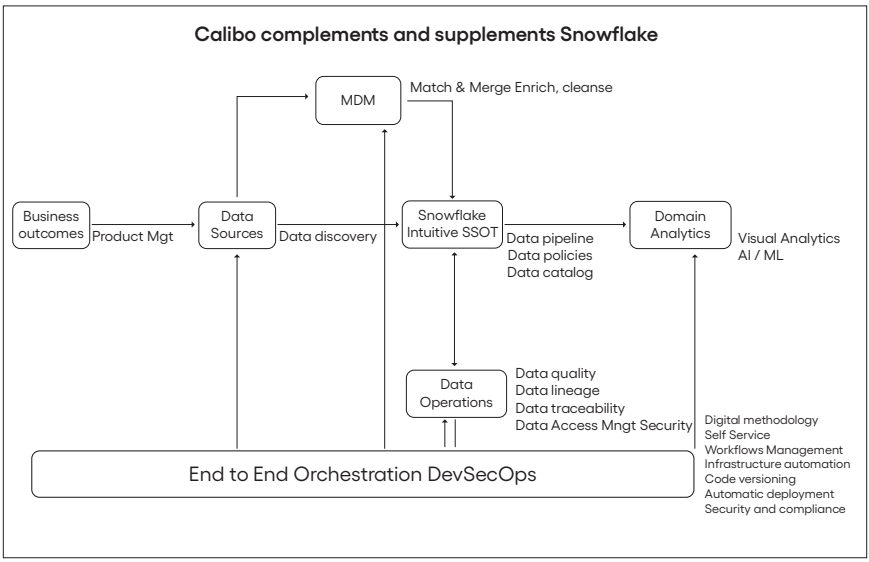


Calibo partners, complements, and supplements Snowflake by bringing capabilities that highlight bringing value to data. By using

> "Every enterprise is pushing to improve business results. Data is one of the key ingredients to new business models and insights in every industry."

Calibo, Snowflake customers can get faster adoption and accelerate the product creation process by 50%.

The diagram below demonstrates Calibo key features to complement Snowflake.



How can this best be achieved? Here are some steps to follow.

Step 1: Business Outcomes

Calibo provides an embedded, yet flexible, digital business methodology that enables companies to fully leverage an Agile, product-centric approach. This methodology helps business resources to work closer with technologists, allowing the business to take more ownership. It simplifies and brings consistency and visibility to digital product creation.

It allows business to democratize value to data. Business engages and owns core data assets in a cloud agnostic environment using self-service and automated features.

Calibo assists product teams in managing their product tasks. It simplifies incremental delivery using a bite-size approach enabled by Agile processes. The monitoring and metrics allow organizations to measure the efficiency and productivity of teams as well as maturity over a period of time. Calibo also provides the ability to define policies, so engineers conform to defined rules such as which technology or machines can be used.

Step 2: Data Source and Data Discovery

Calibo uses data crawlers to gauge data sources and map data elements such as meta data into data catalogs. The crawlers assist in data discovery and help users to understand relationships between data elements. Automatic data crawlers will capture meta data and make it visible through a simple user interface so users can view and decide how to make use of them.

The self-data discovery requires same data to be accessible in data catalogs. The catalogs will help business users to understand existing data sets, ownership, treatment, and access.

Calibo allows the integration of metadata into key data catalogs, such as AWS Glue and soon in Alation and Collibra.

Step 3: Intuitive SSOT

A lot of organizations have fragmented and heterogeneous data sources. Data needs to be discovered, integrated, enriched, and made available for consumers—this is SSOT. They need to integrate data sets to create the SSOT. This is done by enriching data sets and storing them into a common data store.

SSOT can be either a data lake or a data mesh. A data lake provides common data standards into a unique data architecture for all an organization's business domains. This is a one size fits all approach.

Data mesh provides a domain-driven architecture and the ability to differentiate based on domain needs.

In both cases, teams need to create a unique, enriched data model that can be used to serve the enterprise.

Calibo offers the ability to create data meshes and data lakes. Organizations can use its policies and standards template to define a mesh or lake architecture. Customers define the technologies and methods they want to use for enterprise or for domain.

The data intelligence studio allows users to discover data through automated crawlers, integrate into data catalogs and process data into a variety of steps using a simple wizard. They will see data steps such as ingestion, transformation, and quality that exhibit Snowflake capabilities.

Step 4: MDM

The creation of SSOT requires the ability to manage data and create a unique referential that people trust and use. This is MDM. Master data management helps to define unique data entities that can be used and reused and processed to keep high data trust.

These entities must be stored and made accessible to others to consume through APIs. Business stewards are responsible for maintaining them and grant access for others to consume them.

Calibo helps to manage master data. It maps and stores master data and makes it accessible for business users to see through data catalogs and integrate master data in data pipelines through APIs. Calibo native features allow users to match, merge, cleanse and enrich master data so it can be used in analytical data processing scenarios.

In addition, Calibo will integrate to standard MDM platforms, so users can maintain a unique data referential.

Step 5: Data Operations and Governance

In addition to master data, organizations need to look for data operations and governance activities. This will include data quality, data lineage, data security and FinOps.

For data quality, business rules are programmed to ensure data is complete, valid, or enriched. Calibo complements and supplements Snowflake to provide an integrated data quality engine that masters business rules. These rules can be determined in a simple user interface.

With the abundance of data and technologies, it is important to monitor consumption, so storage, compute and license costs are maintained within an organization's limit. Product owners want to know how much it costs to build and operate a product so they can actively adjust.

Calibo offers embedded FinOps capabilities. It allows to monitor consumption of product irrespective of technologies used. Its native dashboard allows to drill down within product and technologies. You can analyze efficiency of code but also to parameter usage (e.g., number of instances, or compute capacity available).

Step 6: Self-Service Domain Data and Analytics

Once data is present in a SSOT and governed through MDM and DataOps, users can focus on creating relevant analytics.

Users will create a refined data layer as a subset of SSOT. This reduced data set will primarily focus on analytics use cases, to optimize run-time performance. All KPI calculations will be processed and stored. The user will create visualizations to access these data sets.

Calibo eases the creation of domain analytics. The data intelligence studio helps users create the data pipelines using a visual graphic interface. It allows users to select data elements to ingest, interpret and integrate into the refined layer.

Calibo also integrates naturally with visual analytics solutions such as Tableau, Qlik or Snowflake Streamlit. Calibo makes the Snowflake data layer automatically available in these tools. Users can launch visualization technology from Calibo and start creating charts and tables.

In advanced domain analytics, users can create AI models and run MLOps to ensure the model remains persistent and trained. Calibo supports the creation of AI models in most modern technologies such as R and Python. It embeds the AI models as a step in the data flow.

In summary, Calibo brings self-service capabilities so any user or product team can build data and analytics products using self-guided methods with ease.

Step 7: Cloud Data Engineering Platform and End-to-End Orchestration

A digital data and analytics product requires access to a development environment with adequate technologies. It also requires the ability to create a code repository and promote code through various stages of development.

This process is tedious and often manual in large organizations, creating points of failure and reducing engineering velocity. Calibo enhances the engineer experience, providing a self-service cloud agnostic environment where engineers can instantiate the environment automatically, deploy the required technologies, and create a code repository with the click of a button.

> "A lot of organizations have fragmented and heterogeneous data sources. Data needs to be discovered, integrated, enriched, and made available for consumers—this is SSOT."

In addition, it provides workflows to guide product teams in the creation process. Calibo, leveraging Snowflake and other industry leading technologies, provides an end-to-end data engineering capability that automates and orchestrates the software and data pipeline development increasing speed, consistency, and scale.

In many enterprises, data trust and security are limiting elements. Data modification is abstract, and people want to know who accessed what information and how it is used.

The fragmented digital ecosystem makes information traceability very difficult. Organizations need to ensure they can map data from sources to target, including all transformation steps. They also need to comply with regulations so information can be reproduced when required.

Calibo is unique as it provides the ability to map an end-to-end data process irrespective of technologies and make it accessible for people to view and drill down.

Instead of defining policies that must be followed manually, Calibo provides the automated guard rails of approved technologies, and required security processes and tools. It provides a versioning audit log and traceability that regulators want to see.

At the same time, organizations want to make data accessible. They know that data democratization requires this, although they want to understand who uses information and for what purpose.

Calibo provides native capabilities to restrict data access to users, groups of users based on roles, and attributes controls. In addition, it integrates to solutions such as Immuta so organizations can trace access of data elements throughout the lifecycle of their data.

Using the Digital Business Methodology and leveraging Calibo as an end-to-end digital ecosystem, organizations can create a holistic approach, create digital products that deliver business outcomes, and drive adoption of Snowflake in less than six months.

Case Study: How a Pharma Data Organization Leverages Calibo to Drive Faster Snowflake Adoption and Value to Data

Background

The firm had embarked on a multi-year journey to simplify and streamline the ingestion, transformation, and visualization of all data and empower business users to create insight and accelerate innovation.

It manages a portfolio of 500 digital products created by more than 100 product teams and provides a modern platform including data engineering and DataOps pipelines as a service to its partner ecosystem.

Overview of the Company

A leading multi-national pharmaceutical corporation focused on innovative science and digital technologies to create transformational treatments in areas of great medical need.

Targeted Business Outcomes

- Reduced time from ideation to product realization from six months to three months.
- Improved engineer productivity by 40%.
- Enhanced IT resource usage and focus on value-add task adopting new age cloud infrastructure and engineering practices.

Approach

A focus on improving on the transformational tenets such as:

- Defining the business value of the transformation.
- Preparing for culture change.
- Starting small but being strategic.
- Mapping out the technology platform approach using Calibo and Snowflake as core.
- Seeking out partners and expertise (four different divisions were considered).
- Gathering feedback and refining as needed.
- Scaling and Transforming.

Activities

- Establish Calibo as the end-to-end ecosystem and self-service development environment.
- Integrate Calibo to the customer cloud technology environment.
- Configure Digital Business Methodology and guardrails in Calibo to enforce consistency.
- Create SSOT and digital products per domain leveraging Snowflake as a data mesh architecture.
- Automate and orchestrate DevSecOps to limit manual tasks and hand offs between teams and drive efficiency.
- Provide business ownership with their data using data operations features of Calibo.

Outcomes Achieved

- 50% reduction in product creation from ideation to MVP.
- 30% engineering productivity.
- 15% reduction in Cloud consumption.

Case Study: How a Digital Media Provider Can Build New Digital Services Leveraging Calibo and Snowflake

Background

A leading media company has invested heavily in a modern data strategy leveraging the cloud for data, AI/ML, and analytics. The data strategy was revised to use a hybrid cloud architecture with a focus on using Snowflake as the data core. The new environment should enable the creation of digital product on demand.

Problem Statement

- Platform adoption is low among business users.
- Team is unable to switch from legacy environment, therefore incurring additional costs.
- Lack of visibility of engineering maturity, inadequate tool chain design patterns, and cross-team dependency management impacting quality and speed of scrum teams.

Targeted Business Outcomes

- Improve data accessibility and business ownership.
- Accelerate digital product creation.
- Create standardized engineering environments to avoid inconsistency and rationalize legacy.

Approach

- Establish Calibo as an end-to-end ecosystem and integrate into the customer cloud technology environment.
- Configure Digital Business Methodology and guardrails in Calibo to enforce consistency.
- Create SSOT and digital products per domain leveraging Snowflake as data mesh architecture.
- Establish a migration factory to accelerate data migration from legacy systems into Snowflake leveraging Calibo.
- Rationalize technologies to simplify workflows.
- Provide business ownership with their data using data operations features of Calibo.
- Automate and orchestrate DevSecOps to limit manual tasks and hand offs between teams and drive efficiency.

Outcomes Achieved

- Centralized data store with data products accessible for Digital business
- Improved maturity in Cloud platform engineering with automated governance.
- 30% accelerated data migration into Snowflake.

Calibo

Ready to accelerate your digital journey? Learn more how Calibo can partner with you.