

Deloitte.

In association with **Dataiku**

FASTER, BETTER DATA SCIENCE FOR ALL!



Accelerating data science analyses using industrialized assets with Deloitte's Asset Store

Organizations have long struggled to standardize and expedite data science analyses, and the rise of artificial intelligence (AI) and machine learning (ML) exasperated the burden as organizations struggle with prolonged data preparation cycles, inconsistent analysis methods, and varying team competencies. This means longer wait times to deliver usable data, inconsistent results, and poor AI/ML insights.

It's time to harness the power of industrialized AI to drive data insights with speed and scale. Deloitte's Asset Store with Dataiku and Snowflake enables reusing proven analytics frameworks and data assets, helping teams rapidly scale insights across brands and markets to drive business adoption.

By empowering teams to collaboratively build on trusted assets, the Asset Store democratizes AI throughout the organization, propelling continuous improvement of data-driven strategies and business outcomes.

PRESENT STATE OF DATA ANALYSES: SLOW, INCONSISTENT, INCOMPLETE

Working smarter, not harder, doesn't happen overnight as the path to accelerated data science is riddled with challenges:



LONG PREPARATION CYCLES (MONTHS, NOT DAYS):

Most new data science analyses take months, even for analyses that are well-known and undertaken previously for other brands or markets.



INCONSISTENT ANALYSIS APPROACHES:

Data science projects often lack standardization, as different teams build their own version of the same analysis.



VARYING TEAM COMPETENCIES:

Varied skill levels across teams can hinder the adaptation of complex analyses across brands/markets, raising questions about trustworthiness of insights.



LACK OF DATA DISCOVERABILITY:

Data lives in isolated silos, which can lead to incomplete insights and missed opportunities.

These challenges make it difficult to deliver data science insights on time and at scale, which leads to inconsistent results, increased errors, and no proven track record from which to build trust. For instance, a marketing segmentation analysis in one market might showcase industry-leading insights, while a parallel effort in a different segment could yield subpar results.

Additionally, when only a handful of data scientists are called upon for analyses, these few individuals can easily become overworked, resulting in longer wait times or missed opportunities altogether. This leads teams to continue relying on traditional analytics and resist smarter, data science generated insights.

Streamlining data preparation is key—this laborious task typically consumes¹

80%

of a data scientists time

A BETTER WAY FORWARD

The Asset Store combines Deloitte's depth and breadth of industry experience with the integrated technical platforms of Dataiku and Snowflake to help organizations easily build AI and Data-as-a-Service (DaaS) into their processes and operations.

With a library of pre-engineered feature sets and parameterized analyses templates, the Asset Store gives data scientists and engineers an evolving advantage when tackling similar use cases—this shortens the journey from data to insights and enhances knowledge management.

These Asset Store capabilities help data scientists and other users enhance usability and efficiency ➔



SEARCH AND DISCOVERY

- Data scientists can quickly search and find relevant datasets and analysis pipelines
- Use tags such as “market” or “use case”



REUSE AND ADOPTION

- Data scientists can customize published feature sets and logic for their own use cases
- Reuse feature sets – allows selection of specific versions, snapshots, sub-features
- Reuse analysis pipelines – allows re-configuration using a new dataset and parameters



EASE OF PUBLISHING

- Self-service publication of highly reusable, foundational datasets
- Use custom tags such as “brand” or “key business question (KBQ)” for better discovery
- Publish feature sets to Dataiku Feature Store



FAST EXECUTION

- Quickly execute imported logic pipelines in new projects by pushing down processing to Snowflake – run pipelines in minutes, not hours
- Easily import by sharing references to Snowflake tables rather than recreating tables



VERSION FEATURE SETS

- Version feature sets to recognize logical updates. Users can select specific version number when importing
- Track metadata in a common database table supporting the Asset Store across projects



USAGE TRACKING

- Track how many times feature sets were used across projects

PROGRESS IN MOTION: THE WORKFLOW IN ASSET STORE

Three components support publishing and re-using feature sets and logic pipelines in the Asset Store: Publisher, App Designer, and Asset Store web application.

Component 1 Publishing feature sets to Asset Store

The custom “Publisher” plug-in from Dataiku DSS (Data Science Studio) enables data scientists to publish reusable datasets to the Asset Store with rich metadata that facilitates discovery for other analyses.

A global life sciences and biopharmaceutical provider that leverages the Asset Store’s reusable assets feature expects to reduce most analysis timelines from 12 weeks down to 6 weeks.

50%
faster analyses

Component 2

Discovering and using feature sets/logic pipelines

The Asset Store web application works as a single “storefront” for all reusable datasets and parameterized pipelines to perform various analyses. It enables easy asset discovery and “pulling” the selected asset into the user’s project workflow (Figure 1).

The screenshot shows the Asset Store interface for a dataset named 'vinda_specific_comorbidity'. The interface includes a search bar, a filter dropdown, a table of assets, and an 'About' panel on the right. The table lists the asset name and description, with buttons for 'Import Dataset', 'Clone Logic', 'Historical View', and 'Additional Info'. The 'About' panel provides details such as 'Features', 'Usage Details', 'Source Project', 'Publisher', and 'Tags'.

1. Use the “Filter” and “Group by” functionality to explore assets.
2. Click “Import Dataset” to use the latest version of data in the project.
3. Use “Clone Logic” to get to the App Designer instance for a particular parameterized pipeline.
4. Use “Historical View” to select an older version or snapshot of the data.
5. Use “Additional Info” to bring up a lineage pane, which can be used to review a sequence of recipes and code
6. Use the “About” pane to review available features

A global life sciences and biopharmaceutical provider that uses the Asset Store projects roughly 385% ROI per year from saved data science hours.

385%
return on investment

Component 3 *Reusing logic pipelines (configuration)*

An Industrialization team will parameterize an analysis pipeline and create an app using Dataiku App Designer. When users click “Clone Logic,” the Asset Store creates a new instance of the app designer project, ready for configuration as needed.

End-to-end analytics marketplace

The Asset Store serves as a marketplace to discover and reuse multiple types of analytics assets, including:

Analytic data assets: Standardized, discoverable, and analytics-ready data sets enriched with analytical features.

Parameterized pipelines: Generalized modules or end-to-end analyses that represent building blocks and can be re-configured to adapt an analysis from one brand/market to another.

READY TO GET STARTED?

By integrating AI and machine learning into your organization’s decision-making processes and operational workflows, you’ll drive data insights with speed and scale to transform productivity, and empower data scientists to pursue more innovative analyses in the future.

Please reach out to the Deloitte team to discuss how Dataiku and Snowflake can help you accelerate data science analyses using industrialized assets in the Asset Store.

Steven Zhang

Managing Director
Deloitte Consulting LLP
shenzhang@deloitte.com

Manoj Suvarna

Managing Director
AI & Data Ecosystems
Deloitte Consulting LLP
manoj-suvarna@deloitte.com

Sources

1. <https://techcrunch.com/2021/08/19/companies-betting-on-data-must-value-people-as-much-as-ai/>

About Deloitte

This communication contains general information only, and none of Deloitte Touche Tohmatsu Limited, its member firms or their related entities (collectively, the “Deloitte Network”), is, by means of this communication, rendering professional advice or services. Before making any decisions or taking any action that may affect your finances, or your business, you should consult a qualified professional adviser. No entity in the Deloitte Network shall be responsible for any loss whatsoever sustained by any person who relies on this communication. As used in this document, “Deloitte” means Deloitte Consulting LLP, a subsidiary of Deloitte LLP. Please see www.deloitte.com/us/about for a detailed description of our legal structure. Certain services may not be available to attest clients under the rules and regulations of public accounting.

Copyright 2024 Deloitte Development LLC. All rights reserved.