MyDataWork: Addressing the Last-Mile Governance Gap in Analytics Workflows

Organizations have invested heavily in data platforms, BI tools, and orchestration frameworks — yet a persistent governance gap remains at the point of actual analytical work. MyDataWork is an early-stage SaaS product that directly addresses this gap by giving individual analysts a lightweight workspace to catalog their file and cloud-based assets, document use cases, track measurable objectives, and surface lineage — without requiring enterprise-scale implementation. Early Access users should evaluate it as a practitioner-layer complement to existing data stacks, not a replacement for any category within them.

Market Context

The data and analytics tool landscape has matured considerably across infrastructure, integration, and visualization layers. Modern data stacks now routinely include cloud data warehouses, transformation frameworks, orchestration engines, and self-service BI platforms. What remains consistently underdeveloped is the human context layer — the documentation of why analytical assets exist, who depends on them, how they connect to business outcomes, and what an analyst is actually working on at any given time.

This gap is not a failure of technology adoption. It is structural. Enterprise data catalogs (Collibra, Alation, Atlan) address metadata governance at the platform level, but require significant IT investment, data engineering integration, and organizational change management. They are designed for platform owners, not practitioners. The individual analyst — the person actually building the Excel model, writing the SQL query, maintaining the Power BI dashboard, or managing dbt models in GitHub — has no equivalent tool designed for their workflow.

The result is predictable: institutional knowledge lives in individuals, onboarding is slow, stakeholder communication is reactive, and the business value of analytical work remains largely invisible to leadership until something breaks.

What MyDataWork Does

MyDataWork operates at the practitioner layer. It provides two complementary paths for asset discovery:

- Local file indexing deploys a lightweight Windows-based connector that indexes file metadata — names, paths, types, modification dates, folder relationships — without accessing file contents. Cloud-synced storage (OneDrive, Google Drive,

Dropbox, Box) is fully supported, as these services sync to local folders that the connector scans automatically.

- Cloud source connectors enable direct API-based integration with four platforms — GitHub, dbt Cloud, Databricks, and Snowflake — requiring only an API token and no software installation. GitHub repositories are scanned for SQL, Python, notebook, and other analytical file types. dbt Cloud projects expose models, sources, and exposures. Databricks workspaces surface notebooks and jobs. Snowflake databases yield table and view catalogs with schema and row count metadata.

Assets from both paths appear in a unified workspace where analysts can:

- Catalog assets across tools including Excel, SQL, Python, Power BI, Tableau, Alteryx, CSV, GitHub-hosted code, dbt models, Databricks notebooks, and Snowflake tables
- Document use cases — connecting assets to business objectives, stakeholders, and outcomes
- Track measurable progress — setting baselines, current values, and targets with automatic progress bar calculation and a free-form progress log
- Visualize lineage — automatically inferring relationships between assets based on metadata patterns
- Generate AI recommendations — surfacing insights about asset relationships and use case optimization
- Export portfolios — producing shareable summaries of analytical work for stakeholder consumption

The deployment model is deliberately lightweight: no IT project, no data engineering dependency, no API integration with existing platforms required beyond individual API tokens. An analyst installs the connector or enters cloud credentials, and has a working catalog within minutes.

Strategic Relevance

Complementary, not competitive

MyDataWork does not compete with Snowflake, dbt, Tableau, GitHub, or enterprise data catalogs. It occupies a distinct layer — practitioner-level context management — that none of these tools address at the individual user level. Notably, MyDataWork integrates with several of these platforms as data sources while providing the organizational and outcome context that the platforms themselves do not capture. Organizations already running mature data stacks will find MyDataWork additive rather than duplicative.

Bridges local and cloud analytical environments

The combination of file-based and API-based connectors means MyDataWork serves analysts regardless of where their work product lives — on local or synced drives, in Git repositories, in managed cloud notebooks, or in database catalogs. This is a meaningful differentiator from file-only tools and positions the product for the increasingly hybrid analytical environments that characterize mid-market organizations.

Addresses a documented governance failure mode

Gartner research consistently identifies knowledge loss during personnel transitions, poor stakeholder communication from analytics teams, and the inability to demonstrate ROI from data investments as top challenges for data and analytics leaders. MyDataWork's use case tracking, objective setting, and stakeholder linking directly targets these failure modes with a structured but lightweight approach.

Goal tracking at the practitioner level

The Objectives & Progress capability — baseline, current value, target, unit of measure, and automatic progress calculation — gives analysts a structured way to connect their work to measurable outcomes. This addresses a capability gap that exists across project management tools (which lack analytical context) and data catalogs (which lack outcome tracking).

Scales bottom-up

Unlike top-down governance initiatives that require organizational mandates to succeed, MyDataWork delivers individual value first — making adoption self-motivating. When analysts use it for their own productivity, the team-level and leadership-level visibility benefits emerge as a natural byproduct. This bottom-up adoption pattern is consistent with successful category-defining tools including Slack, Notion, and Figma.

Relevant to the "AI-ready data organization" narrative

As organizations accelerate AI adoption, the ability to understand what data assets exist, how they connect, and who owns them becomes prerequisite infrastructure. MyDataWork builds this understanding at the practitioner level, where it is most accurate and most current.

Limitations and Considerations

Windows-only file connector

The local file indexing connector is Windows-exclusive. Analysts whose work product lives entirely on Mac or Linux will need to rely on cloud source connectors for asset discovery, which covers GitHub, dbt Cloud, Databricks, and Snowflake but not arbitrary local file systems.

Cloud sources in early access

The API-based cloud connectors (GitHub, dbt, Databricks, Snowflake) are available in the current release but are newly introduced. Organizations evaluating these integrations should conduct hands-on testing with their specific environments and credential configurations.

Early stage

MyDataWork is in early access. Prospective users should evaluate it as an emerging tool with the corresponding expectations around feature completeness, stability, and vendor maturity.

Not a replacement for enterprise data catalogs

Organizations requiring platform-level lineage, regulatory compliance metadata, or enterprise-scale governance should continue to evaluate and implement purpose-built catalog solutions. MyDataWork is complementary, not substitutive.

Recommended Actions

For Data & Analytics Leaders

Evaluate MyDataWork as a low-cost, low-risk complement to existing governance initiatives. The practitioner-layer gap it addresses is real and well-documented. The combination of file-based and cloud-based asset discovery makes it relevant across a wider range of analytical environments than file-only tools. Early access pricing removes financial risk from evaluation.

For Analytics Team Leads

Pilot with a small team of analysts as a use case documentation, goal tracking, and stakeholder communication tool. Assess whether the portfolio export function improves leadership visibility into team output, and whether the cloud source connectors surface assets from GitHub or dbt that are not currently visible in any catalog.

For Chief Data Officers

Monitor as an emerging category. The practitioner-layer context management space — combining asset cataloging, outcome tracking, and cloud source integration in a lightweight deployment model — has no clear incumbent. If the bottom-up adoption pattern holds, tools in this category may become standard complements to enterprise data catalogs within two to three planning cycles.

Competitive Positioning

Dimension	Enterprise Data Catalogs	BI / Analytics Platforms	MyDataWork
Target user	Platform / IT teams	Business analysts, executives	Individual data practitioners
Deployment	Enterprise IT project	IT-managed or self-service	Self-install, minutes
Asset scope	Platform-level metadata	Data visualization & reporting	Files + GitHub + dbt + Databricks + Snowflake
Outcome tracking	None	None	Baseline / current / target + progress
Governance model	Top-down mandate	Tool-specific	Bottom-up adoption
Investment required	High	Medium-High	Low
Time to value	Months	Weeks	Minutes

Data catalogs govern enterprise metadata, BI tools report on data. MyDataWork helps individual practitioners organize the working context behind real data work.

Summary Assessment

MyDataWork addresses a genuine and persistent gap in the analytics governance stack. Its practitioner-first design, dual-path asset discovery (local files and cloud API sources), structured outcome tracking, and bottom-up adoption pattern differentiate it meaningfully from enterprise catalog solutions and file-only tools alike. For organizations seeking to improve analytical knowledge management, stakeholder communication, and the visibility of data work without a large implementation investment, MyDataWork merits evaluation.