MyDataWork: The Work Context Layer for AI-Era Analytics Organizations

Organizations have invested heavily in data platforms, BI tools, and orchestration frameworks — yet a persistent gap remains at the point of actual analytical work. MyDataWork is a SaaS product that directly addresses this gap by giving individual analysts and teams 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. As enterprise AI adoption accelerates, the practitioner-layer context that MyDataWork captures is increasingly recognized as prerequisite infrastructure for AI systems to operate meaningfully on real organizational work.

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 work 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.

The conversation around enterprise AI has crystallized this gap in newly relevant terms. Industry discussion through 2025 and into 2026 has shifted from "do we have enough AI capability" to "do we have the context to use AI effectively." Data platform vendors are responding with ontology-graph products designed to give AI semantic understanding of enterprise data. But data context is only part of the picture. AI systems operating in enterprises also need work context — what analysts are building, who owns each artifact, what business outcomes the work serves, what depends on what. This second layer of

context is what MyDataWork captures, complementary to data ontology efforts at the platform layer.

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, creation and 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. The connector performs shallow content inspection on specific file types to extract structural metadata: SQL table references from .sql files, sheet names from .xlsx, column names from .csv, embedded SQL from .ipynb notebooks, connection hints from .yxmd Alteryx workflows, object kind / tables / connections from ThoughtSpot .tml files, project structure from Dataiku .zip exports, and dimensions / measures / sql_table_name references from Looker .lkml files.

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, Jupyter notebooks, Power BI, Tableau, Alteryx, CSV, ThoughtSpot, Dataiku, Looker LookML, GitHub-hosted code, dbt models, Databricks notebooks, and Snowflake tables
- **Document use cases** with structured fields including objective, target outcome, baseline / current / target measurements with automatic progress calculation, priorities, effort estimates, target dates, free-form progress notes, communication logs with stakeholders, and linked action plans
- **Track measurable progress** with quantitative outcome tracking that connects analytical work to business results
- **Link stakeholders** to use cases, capturing who depends on each piece of analytical work

- **Visualize lineage** through automatic inference of relationships between assets based on metadata patterns, with manual augmentation available
- **Engage AI assistance** through multiple integrated AI capabilities: per-asset and per-use-case recommendations, cross-portfolio leverage analysis, agentic opportunity identification, marketplace and migration recommendations, and an AI chat assistant grounded in workspace context
- **Use the Workspace Agent** for on-demand workspace review across cleanup, activity, and insight observations — a hybrid system where deterministic rule-based detection identifies findings and a language model handles human-readable phrasing of what was found
- **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.

Team Privacy Model

MyDataWork's Team plans implement a private-by-default architecture that distinguishes it from shared-workspace tools where all collaborators see everything.

Each Team member works in their own private view of their assets and use cases. Other members cannot see another's work by default. Members selectively share specific assets to a team bulletin board via an explicit Share to Team action; shared assets appear to all team members in a dedicated Shared view. Team admins see the same default view as members, with explicit admin-only opt-in to view the full workspace catalog when needed for onboarding, audit, or oversight tasks. Aggregate metrics — total assets, sharing rates, reuse multipliers — are visible to admins without exposing individual members' private work.

This model addresses a specific anxiety analysts have expressed about catalog tools: that documenting their work means exposing drafts, exploratory work, and scratch analyses to their manager or teammates. MyDataWork's privacy-by-default-with-explicit-sharing approach allows analysts to catalog comprehensively without that exposure, increasing the likelihood of complete and honest documentation.

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 work 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 documented governance failure modes

Industry 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.

Outcome tracking at the practitioner level

The Objectives & Progress capability — objective, baseline, current value, target, unit of measure, priority, effort estimate, target date, progress notes, 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.

Provides the work context AI systems need

As organizations accelerate AI adoption, the ability to understand what data assets exist, how they connect, and who owns them becomes prerequisite infrastructure. Data platforms address this at the semantic layer through ontology graphs and data lineage. MyDataWork addresses it at the work layer — capturing the SQL files, notebooks, dashboards, models, and the analysts who own them, the use cases they support, and the business outcomes they serve. The two layers are complementary. AI agents that need to operate meaningfully in an enterprise need both.

The Workspace Agent within MyDataWork demonstrates this principle directly. The agent surfaces observations about cleanup opportunities, activity gaps, and analytical insights by combining deterministic detection logic with language model phrasing — a hybrid pattern increasingly recognized as the credible path for narrow-scope AI agents. The agent doesn't try to be a general-purpose reasoning system; it surfaces what rule-based checks already concluded, in clear language. This design philosophy distinguishes MyDataWork's AI integration from approaches that wrap a language model in a loop and hope for the best.

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. Mac OS connector support is on the product roadmap.

Pre-Series A scale

MyDataWork is a generally-available SaaS product with established paying customers, a 14-day free trial, and four pricing tiers ranging from Solo Monthly through Team Growth. The company is pre-Series A, and prospective customers should evaluate it with corresponding expectations around feature roadmap velocity, vendor maturity, and support model. Enterprise certifications including SOC 2 Type II are on the product roadmap; the underlying AWS infrastructure maintains SOC 2, ISO 27001, PCI DSS, and other certifications.

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 — now spanning thirteen analytical file types and four cloud platforms — makes it relevant across a wider range of analytical environments than file-only tools. The 14-day free trial removes financial risk from evaluation.

For Analytics Team Leads

Pilot with a small team of analysts as a use case documentation, outcome tracking, and stakeholder communication tool. Assess whether the portfolio export function improves leadership visibility into team output, whether the Workspace Agent's findings produce actionable cleanup and insight observations, 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 work context management space — combining asset cataloging, outcome tracking, team privacy controls, and integrated AI capabilities 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, and may become essential infrastructure for organizations deploying AI agents that need work-layer context to operate effectively.

For Investors

Watch for adoption signal in the mid-market analytics segment. The category — practitioner-layer work context for analytics teams — is currently uncrowded and increasingly relevant to the broader AI enterprise narrative. Key metrics to track in vendor evaluation: time-to-first-cataloged-asset for new users, percentage of cataloged assets that get linked to documented use cases, and the rate at which Team plan workspaces accumulate cross-member sharing activity (an indicator that the private-by-default model is driving honest cataloging rather than suppressing collaboration).

Competitive Positioning

Dimension	Enterprise Data Catalogs	BI / Analytics Platforms	MyDataWork
Target user	Platform / IT teams	Business analysts, executives	Individual data practitioners and small teams
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 + ThoughtSpot + Dataiku + Looker LookML
Outcome tracking	None	None	Objective, baseline / current / target, priorities, effort, target dates, action plans, communication logs
Team privacy model	Top-down governance	Tool-specific permissions	Private-by-default with explicit sharing
AI integration	Emerging at platform layer	Embedded assistance in product	Per-asset recommendations, portfolio analysis, Workspace Agent, AI chat with workspace context
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 and small teams organize the work context behind real analytical activity — and increasingly, provides the work-layer context that AI systems need to operate meaningfully in the enterprise.

Summary Assessment

MyDataWork addresses a genuine and persistent gap in the analytics governance stack — the practitioner-level work context layer that has remained underdeveloped while infrastructure, integration, and visualization layers have matured. Its practitioner-first

design, dual-path asset discovery across thirteen file types and four cloud platforms, structured outcome tracking with team-level privacy controls, integrated AI capabilities including the hybrid-design Workspace Agent, 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, the visibility of data work, and the work-layer context that enterprise AI initiatives increasingly require, MyDataWork merits evaluation. The combination of complementary positioning relative to existing data infrastructure, low implementation cost, and direct relevance to the contemporary AI-context conversation suggests a product worth watching as enterprise data and AI strategies converge.