



ARCHITECTURE WHITEPAPER

Building a data architecture for AI. Two paths, one platform.

How to give AI assistants and agents business data they can be trusted with - and how to decide between a turnkey service and a connectivity layer for the data platform you already run.

Published

October 2026

Dataddo



AT A GLANCE

One question decides the path: do you run a data platform?

NO WAREHOUSE? TURNKEY

Data to AI™

Connectivity, storage, governance, lineage and security in one managed service, from source to AI answer.

- Bring only your data sources and your AI client
- Claude, ChatGPT, Gemini, Copilot and more over MCP
- No warehouse, pipelines or engineering project
- Governed, described data for more reliable answers
- Up and running without a build phase

HAVE A PLATFORM? PLUGGABLE

Data Anywhere™

The connectivity layer for your warehouse, lake or lakehouse, with every Data to AI™ capability included.

- ETL, ELT, CDC, reverse ETL and streaming in one platform
- Snowflake, BigQuery, Redshift, Databricks and data lakes
- Agentic data streaming for real-time agents
- Private data plane in your cloud or on premises
- Feeds OpenAI Data agent and Claude Dashboards

Start with either and grow: **Data Anywhere™ includes every Data to AI™ capability.**



Executive summary

The right AI data architecture depends on one question: do you already run, or plan to run, a data warehouse, lake or lakehouse? If not, **Dataddo Data to AI™** gives your AI assistant governed business data with nothing else to build. If yes, **Dataddo Data Anywhere™** is the connectivity layer that keeps that platform, and every AI tool reading from it, complete and current.

AI answers are only as good as the data behind them. In a 2026 survey of nearly 1,300 IT leaders, about 80% said limited data access constrains their AI initiatives, and only 18% said their data is fully governed. The barrier is rarely the model; it is access, context, freshness and governance of the data.

Data to AI™ is the turnkey path. It connects 450+ business sources, stores the data in a query-ready SmartCache, applies PII masking, quality checks, access policies, lineage and an audit trail, and serves it to Claude, ChatGPT, Gemini, Copilot and other clients over the open Model Context Protocol (MCP). There is no warehouse and no pipeline to build.

Data Anywhere™ is the pluggable path. It moves data from 450+ sources into Snowflake, BigQuery, Redshift, Databricks and open table formats through ETL, ELT, CDC, reverse ETL and streaming. It adds agentic data streaming, private data planes in your own cloud or on premises, and enterprise controls such as SSO, RBAC and bring-your-own-key encryption. It also includes every Data to AI™ capability.

If your situation is...	Choose
No warehouse; business teams want answers from SaaS data in Claude or ChatGPT now	Data to AI™
A data team and an existing or planned warehouse, lake or lakehouse	Data Anywhere™
Agents that must act on changes within a second	Data Anywhere™ agentic data streaming
Regulated or sovereign data that must never leave your perimeter	Data Anywhere™ private data plane
OpenAI Data agent or Claude Dashboards on top of your warehouse	Data Anywhere™ feeding the warehouse

1 Why AI needs a data architecture

A large language model knows language and general facts. It does not know your revenue, your customers, your stock levels or what your team means by an "active account". Every answer about your business depends on the data the model is given, and on how that data is delivered.

That is why most AI projects stall on data rather than on models. In Cloudera's 2026 Data Readiness Index, about 80% of nearly 1,300 IT leaders said limited data access constrains their AI and data initiatives, and fewer than one in five said their data was fully governed.

1.1 What AI needs from business data

Whatever tools are used, an architecture that serves AI has to deliver six properties. Missing any one of them shows up as wrong, stale or unsafe answers.

Property	What goes wrong without it	What the architecture must provide
Access	Data sits in dozens of SaaS tools, databases and files the model cannot reach	Reliable connectors to every source, maintained as APIs and schemas change
Context	The model sees cryptic field names and guesses joins, filters and metric definitions	Descriptions, a business glossary, metric definitions and relationships
Freshness	Answers describe yesterday's world; agents act on outdated facts	Sync schedules or change streams matched to how fast decisions are made
Quality	Duplicates, gaps and broken loads produce confident but wrong numbers	Validation, deduplication and anomaly checks before data is used
Security	Personal or confidential data reaches a model or a person it should not	Access policies, PII masking, encryption and control over where data is processed
Lineage and audit	Nobody can explain where a number came from or who asked for what	Lineage from source to answer and a log of every query

Table 1. Six properties of AI-ready data, and what happens without them.

1.2 The layers of a data architecture for AI

Modern data architectures share the same building blocks, whether they are assembled from separate products or delivered as one service. Governance is not a layer of its own; it runs across all of them.

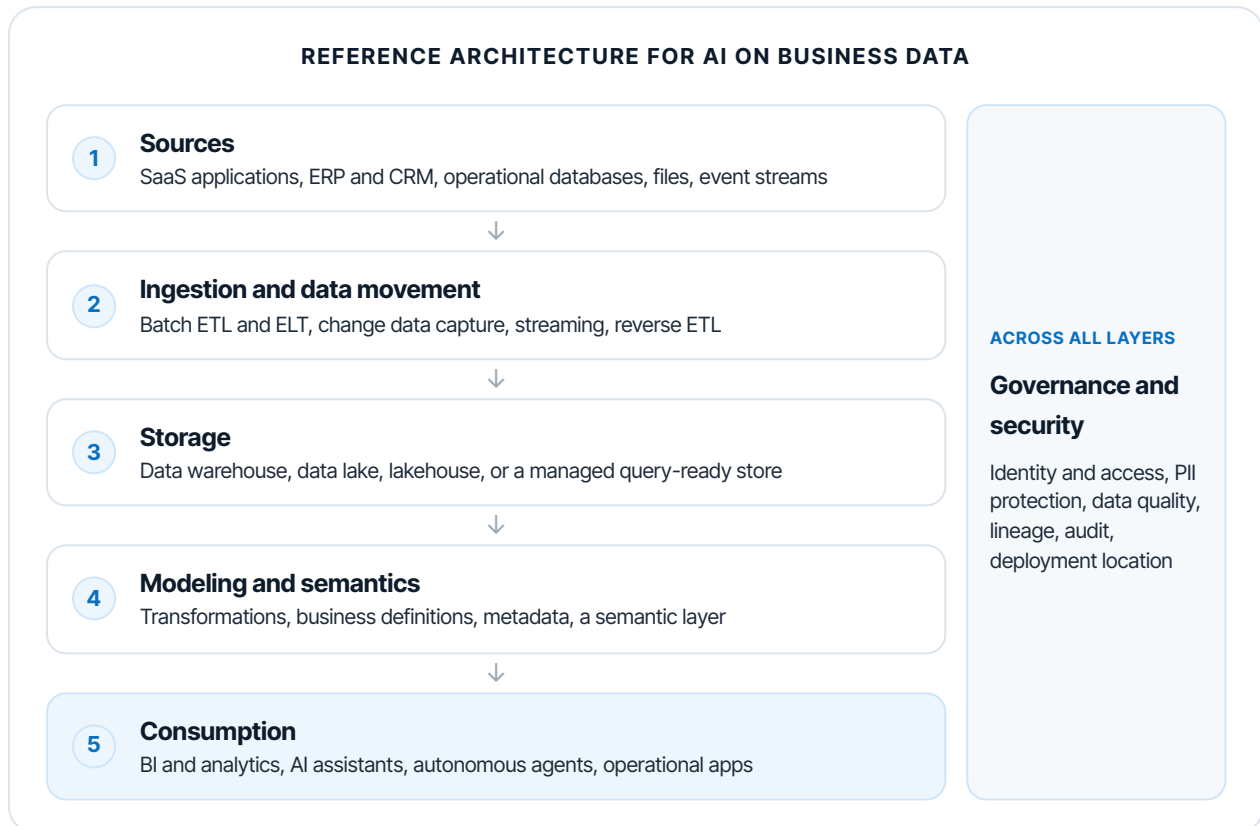


Figure 1. Generic building blocks of an AI data architecture. Each layer can be a separate product or part of one service.

1.3 How AI tools reach the data

AI clients connect to business data in three common ways. Most organisations end up using more than one.

- **Query on demand.** The assistant asks a governed data service for exactly the figures it needs. The Model Context Protocol (MCP) has become the common standard here: an open-source standard for connecting AI applications such as Claude or ChatGPT to external systems, described by its maintainers as "a USB-C port for AI applications".
- **Analysis over a warehouse.** AI analytics tools write and run SQL against a warehouse or lakehouse. They are only as complete and current as the tables they read.
- **Event-driven agents.** Agents that act on operational changes, such as an order, a payment or a support ticket, subscribe to streams of changes rather than waiting for the next batch.

1.4 Batch or real time

Freshness is a business decision, not a technical default. A weekly marketing review is well served by data loaded every few hours. A support agent checking an order, or a fraud agent scoring a payment, needs changes within seconds. Real-time pipelines cost more to run and operate, so the right architecture uses each where it pays off.

1.5 Common failure modes

Failure mode	What it looks like	Root cause
Confident wrong answers	A plausible figure that does not match the data	No context: the model improvises definitions and calculations
Stale answers	Correct logic applied to outdated data	Refresh cadence not matched to the use case
Shadow AI	Employees paste exports into unapproved tools	No sanctioned, easy route from business data to AI
Data exposure	Personal or confidential fields reach a model	Raw access without masking or permissions
Copy sprawl	Agents and teams keep their own extracts that drift apart	No single governed source of truth
Unexplainable numbers	Nobody can trace where an answer came from	No lineage or query audit

Table 2. How AI data projects typically go wrong.

1.6 Build or buy

There are two broad ways to put these layers in place. The first is to assemble them: an ingestion tool, a warehouse or lakehouse, transformation, a catalog, governance tooling and an AI integration, operated by a data team. It offers the most control and suits organisations that already run a data platform. The second is to adopt an integrated service that delivers the layers end to end. It is faster and needs no engineering team, at the cost of less architectural freedom.

Dataddo offers both routes on one platform (Section 2): Data to AI™ is the integrated, turnkey route (Section 3), and Data Anywhere™ is the connectivity layer for the assemble-it-yourself route (Section 4).

2 The Dataddo platform at a glance

Data to AI™ and Data Anywhere™ run on the same platform. It is built around eight ideas: one control plane over many data planes, every transport pattern in one tool, independence from any single vendor, security in the architecture, three interfaces for three audiences, incremental adoption, ownership of the data contract, and native fit with the enterprise systems around it.

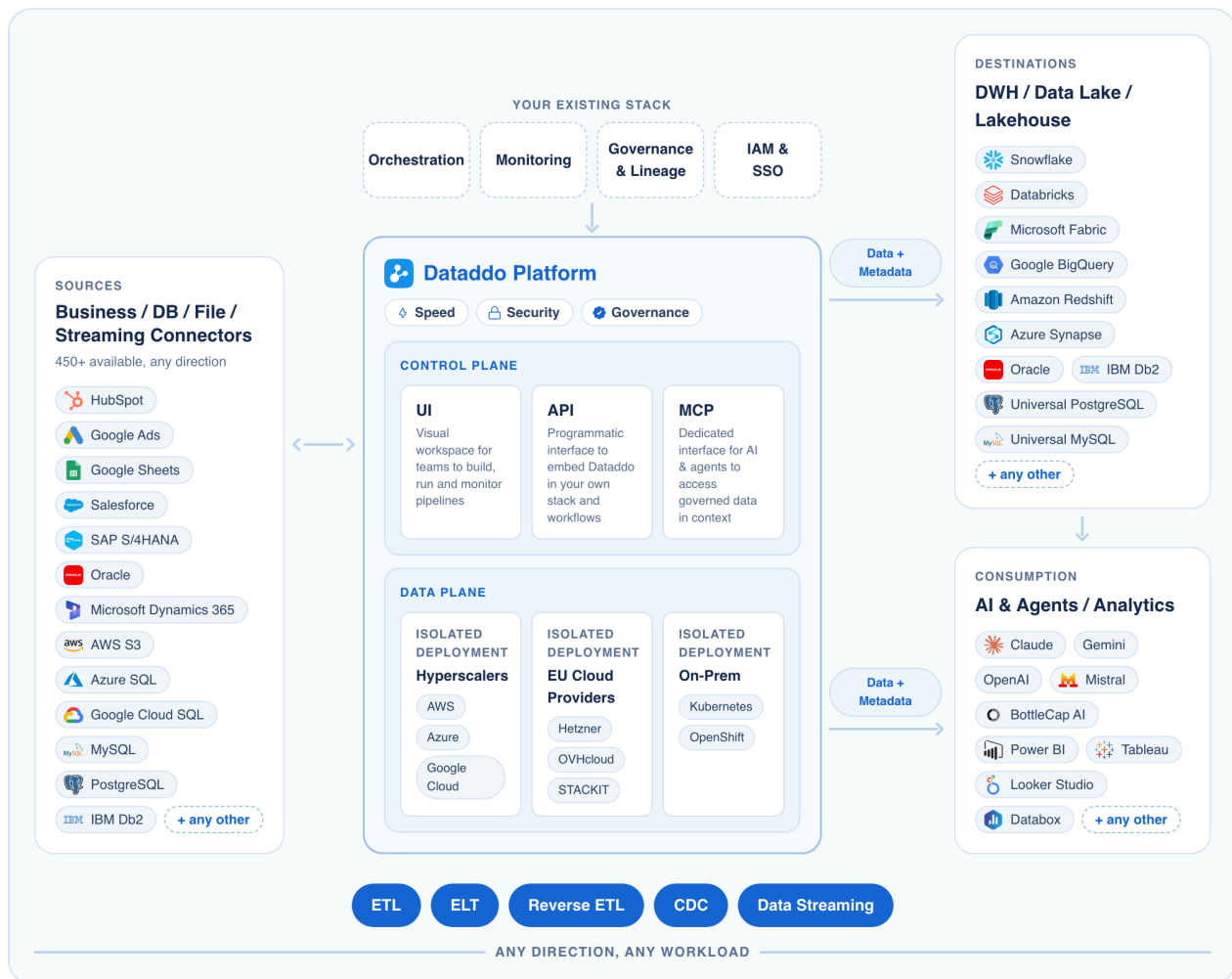


Figure 2. Dataddo platform architecture, as published on dataddo.com/platform.

2.1 Multiple Data Planes. Single Control Plane.

The platform separates deciding from doing. The cloud control plane, managed by Dataddo, handles orchestration, configuration, credentials, governance and monitoring; it never sees or stores your records. Data planes do the extraction, transformation and loading, and run wherever your data lives: in the cloud, on premises, or both.

- **Many data planes, one control plane.** One control plane orchestrates several data plane deployments at once, giving centralized governance without giving up architectural flexibility.

- **Sensitive data stays private.** Sensitive data can be processed only inside a private, on-premises data plane, while less sensitive workloads run in the cloud, all under one management interface.
- **Deploys anywhere.** Data planes run on AWS, Azure and Google Cloud, on sovereign, private and regional clouds, and on on-premises container platforms including Kubernetes, Red Hat OpenShift and VMware Tanzu.

2.2 Every data transport pattern. One platform.

Data moves in different ways for different reasons. The platform supports the full range of transport patterns, so a team never needs a separate tool just because the delivery pattern changed. All patterns share the same connectors, monitoring and access control.

Pattern	What it does	Typical business use
ETL and ELT	Classic extract-transform-load, or load first and transform in the warehouse	BI dashboards and reporting, marketing and sales analytics, centralizing SaaS data in a warehouse
Change data capture	Real-time, low-latency replication that tracks row-level changes as they happen	Real-time warehouse replication, operational reporting, keeping analytics in sync with production
Data streaming	Continuous, event-driven pipelines for time-sensitive and AI-ready workloads	Event-driven pipelines, real-time analytics, keeping AI context and RAG stores fresh
Reverse ETL	Pushes curated, governed records from AI agents or the warehouse back into CRMs and operational systems	Syncing scores and segments to CRMs, activating AI-generated insights
Batch file delivery	Structured delivery of datasets as files to S3, SFTP, Google Cloud Storage, Azure Blob Storage or any storage target	Scheduled exports, partner and vendor data exchange, archival and compliance
Direct connection to AI and agents	Serves governed data straight to LLMs and agents over MCP, with no data warehouse in between	AI answers across all your tools, agentic workflows on live business data

Table 3. Data transport patterns supported by the Dataddo platform.

2.3 Vendor- and stack-agnostic by design - and EU sovereign by choice.

Dataddo is the integration layer above the data stack, not tied to any single vendor. Whether the destination is Snowflake, Databricks, Microsoft Fabric, BigQuery or an open standard such as Apache Iceberg, the platform connects to it. When data residency matters, it runs entirely within EU sovereign clouds.

Category	Supported
Warehouses and lakehouses	Snowflake, Databricks, Microsoft Fabric, Google BigQuery, Amazon Redshift, Azure Synapse, ClickHouse, Firebolt, Vertica and many others
Open table formats	Apache Iceberg, Delta Lake, Apache Hudi
Public clouds	AWS, Azure, Google Cloud
EU sovereign clouds	Hetzner, OVHcloud, STACKIT
On-premises data plane	Kubernetes, OpenShift, VMware Tanzu

Table 4. Where the platform connects and runs.

2.4 Built for high-security, high-sensitivity, high reliability

Security is part of the architecture, not added afterwards. From network topology to field-level data handling, every layer is designed for strict compliance and security requirements.

- **Network isolation.** Sensitive data is processed in a fully isolated data plane in your private cloud or on premises. It never needs to cross public infrastructure, and the control plane orchestrates pipelines without seeing the payload.
- **PII detection and data masking.** Built-in tooling identifies personal data and applies masking, tokenization or redaction at ingestion, before data moves downstream. PII fields can also be excluded from extraction entirely.
- **End-to-end encryption and BYOK.** All data is encrypted in transit and at rest. Customers can bring their own keys via AWS KMS, Azure Key Vault or a hardware security module, so the keys stay under their control.
- **High availability.** Multi-region deployments, automated failover and built-in redundancy keep pipelines running when infrastructure fails.

The platform is certified and compliant with SOC 2 Type II, ISO 27001, GDPR, HIPAA, DORA, CCPA, LGPD and POPIA.

2.5 UI, API, MCP

A business analyst setting up one source and a platform engineer deploying thousands of pipelines need different interfaces. The platform provides three, all backed by the same governance, role-based access control and audit layer.

Interface	Designed for	Best for
UI	Business analysts and data teams who build and monitor pipelines without writing code	Self-service pipeline setup, monitoring and fast proofs of concept
API	Developers and platform teams automating Dataddo programmatically	Embedding Dataddo in internal portals, provisioning pipelines at scale, CI/CD
MCP	AI agents and assistants such as Claude, ChatGPT, Cursor, LangChain and LlamaIndex	Letting agents provision pipelines and query governed data in context, over a native protocol

Table 5. Three interfaces to one platform.

2.6 Start with what you need today. Scale to what you'll need tomorrow.

The platform is designed for incremental adoption: no overnight re-architecture and no large upfront investment before returns appear. Teams start with a handful of pipelines, prove value quickly and expand at their own pace.

- **Incremental adoption.** Deploy what is needed now, with no big-bang migration or re-platforming. When the business switches CRM or adopts a new tool, new pipelines are deployed quickly, so the data layer keeps pace with change.
- **Short time to value.** First pipelines are live in hours, not months, and returns start before the rollout is complete.
- **No ceiling.** The platform that runs an initial proof of concept scales to organization-wide deployment without architectural changes.

2.7 We own the data contract. Not just the connector.

Source APIs change, schemas evolve and vendors deprecate endpoints. In most integration setups, finding and fixing that falls to the customer's team. With Dataddo, it falls to Dataddo.

Event	Without Dataddo: your team owns it	With Dataddo: Dataddo owns it
API or auth change	You discover the breakage and scramble to fix it	The connector is updated and the pipeline restored, often before you notice
Schema drift	Columns change and pipelines break or silently corrupt data	Detected automatically and handled by configurable rules
Endpoint deprecated	You re-engineer the integration	Dataddo owns the update; the data contract holds
Missing connector	You build and maintain a custom integration	Built and maintained by Dataddo under an SLA of about four weeks
Silent degradation	You find out when a report or model run fails	Proactive monitoring catches anomalies and delays first
Debugging	You dig through logs across disconnected tools	Run histories, payload inspection and end-to-end lineage in one place

Table 6. Who owns what when a source changes.

2.8 Interoperability with enterprise systems

A modern data platform cannot exist in isolation. Dataddo integrates natively with the systems around the data infrastructure: identity, governance, monitoring, orchestration, security monitoring and secret management.

Area	Integrations
Identity and access management	SAML 2.0, OIDC, RBAC; Okta, Microsoft Entra ID, OneLogin, Keycloak and more
Lineage and governance	Collibra, OpenMetadata, DataHub, Dawiso and more
Monitoring and observability	Datadog, Prometheus, Grafana and more
Orchestration	Apache Airflow, Prefect, Dagster and more
SIEM	Splunk, Microsoft Sentinel and more
Secret management	HashiCorp Vault, AWS Secrets Manager, Azure Key Vault and more

Table 7. Native integrations with enterprise systems.

3 Path A: Data to AI™, the turnkey path

Data to AI™ is a complete, managed service that takes business data from its source to an AI client's answer. The only things you bring are your data sources and the AI client or LLM you already use. Connectivity, storage, governance, lineage and security are built in.

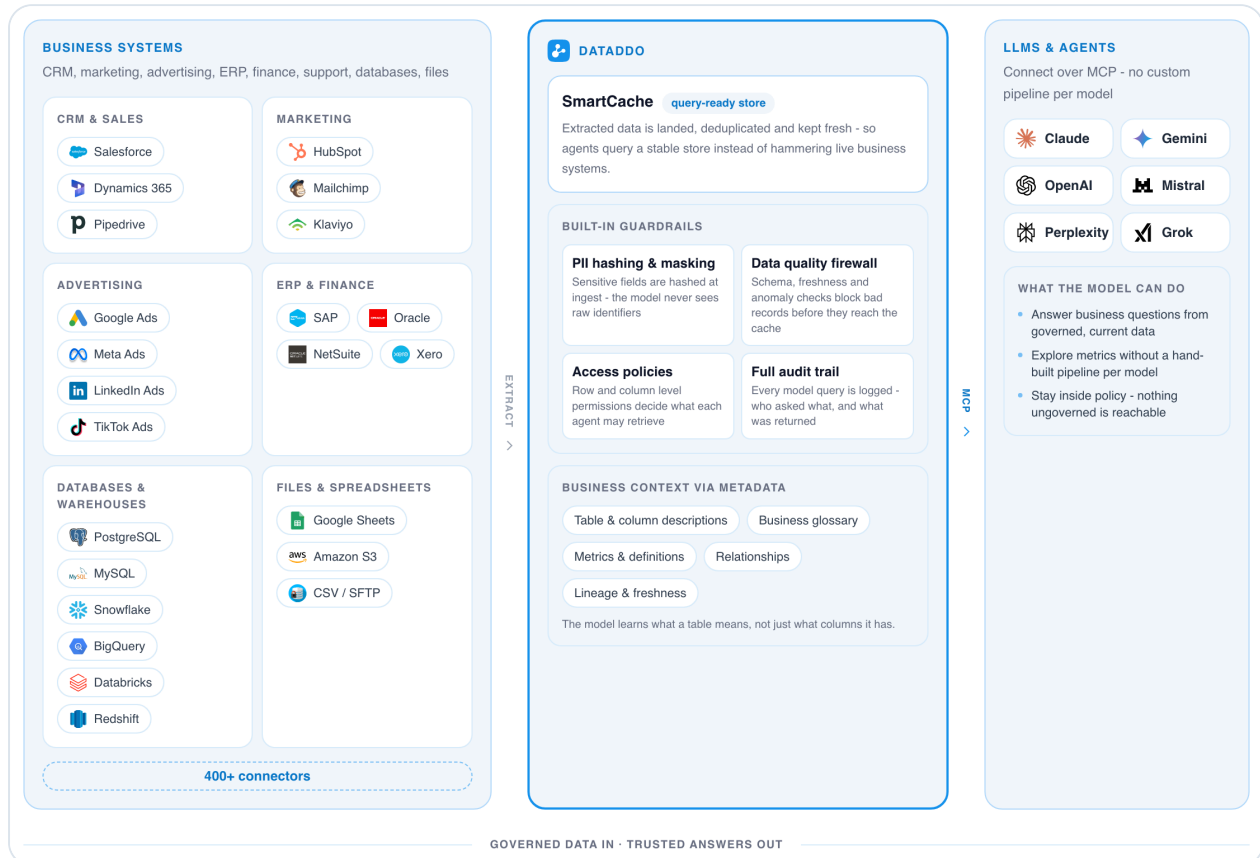
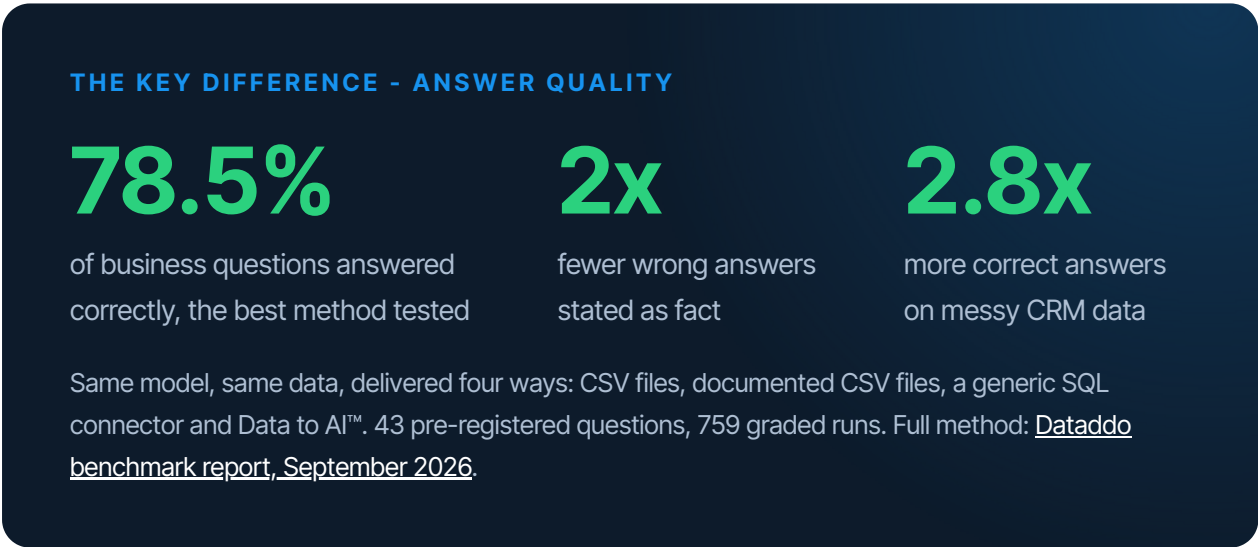


Figure 3. Data to AI™ architecture, as published on dataddo.com/ai-and-agents/claude.

Capability	What Data to AI™ does
Connectivity	450+ managed connectors across CRM, marketing, advertising, ERP and finance, databases and files, maintained by Dataddo
Storage	A query-ready SmartCache keeps extracted data deduplicated and fresh, so no warehouse or data lake is needed
Governance	A data quality firewall runs schema, freshness and anomaly checks and blocks bad records before they reach the model
Security	PII hashed (md5 or xxh3) or masked at ingest; row- and column-level policies decide what each agent may retrieve; read-only access; encryption in transit and at rest
Lineage and context	Descriptions for every table and column, a business glossary, metric definitions, relationships, lineage and freshness
Audit	Every model query is logged: who asked what, and what was returned

Table 8. Everything Data to AI™ includes, so nothing has to be bought separately.



3.1 Deployment

Setup takes three steps: authorize sources from the connector catalog; let Dataddo apply quality checks, PII controls and business metadata; then add Dataddo as a connector in Claude (web, Desktop, mobile or Claude Code) or another MCP client and sign in. If your sources already sync through Dataddo, you add one connection to your AI tool. No SQL, pipeline design or infrastructure work is involved.

3.2 Commercial model

Data to AI™ is priced per seat, like the AI assistants it serves. Because it is turnkey, there is no warehouse to license, no pipeline or transformation tooling to buy, no separate catalog or governance product, and no engineering team needed to build and run them. For organisations without an existing data platform, that removes most of the cost of the build route described in Section 1.6.

3.3 Who it is for

Business teams in marketing, sales, finance and operations who want reliable answers from their SaaS data in Claude, ChatGPT or Copilot this week, and companies without a data engineering team or a warehouse.

4 Path B: Data Anywhere™, the pluggable path

Data Anywhere™ is the connectivity layer for companies that run, or plan to run, their own data platform. It moves data from 450+ sources into any warehouse, lake or lakehouse, and from there into business apps and AI tools, with one control plane over every flow. It includes every Data to AI™ capability. Your data team keeps control of modeling, the semantic layer and which AI tools read the data.

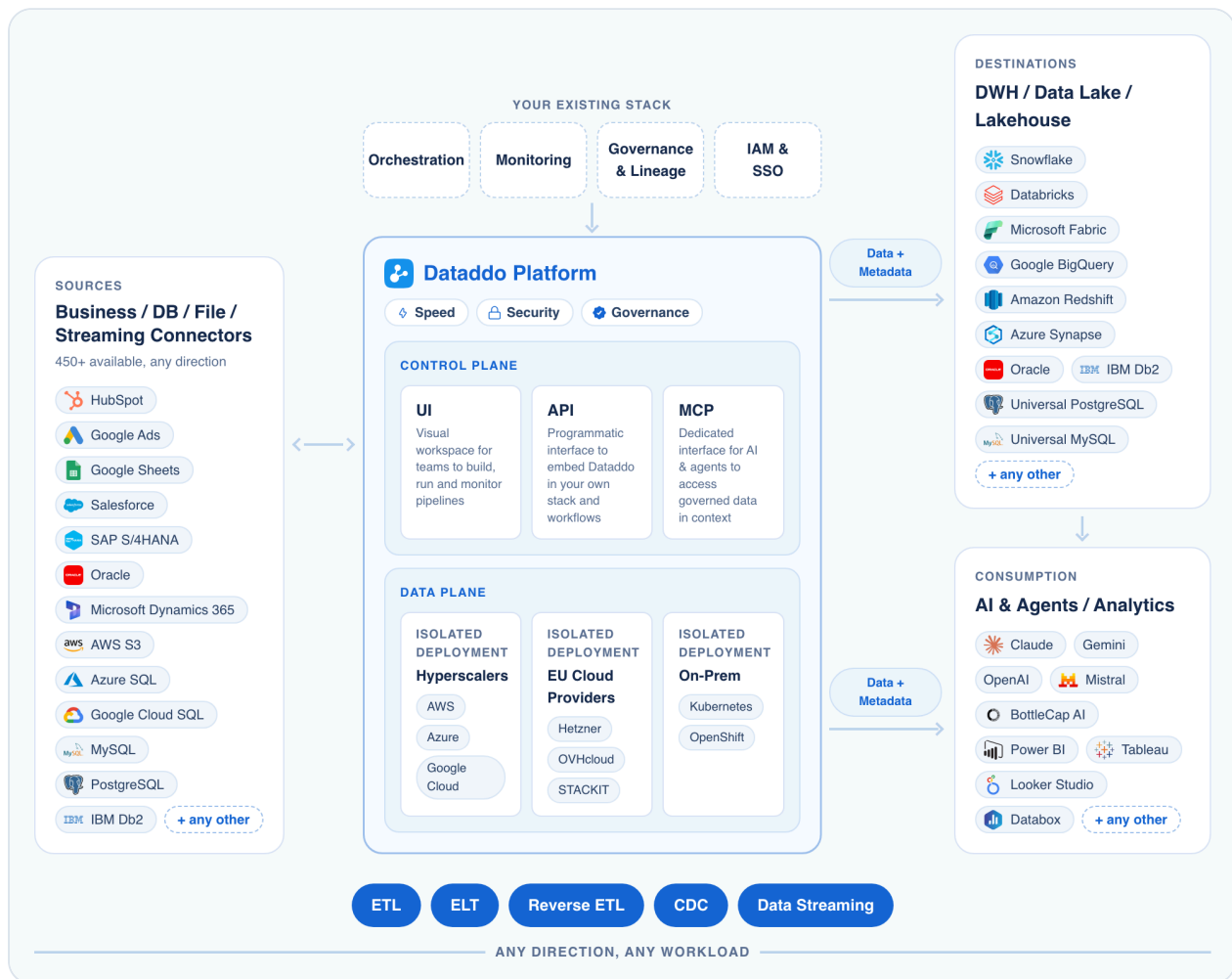


Figure 4. Data Anywhere™ architecture: one control plane, data planes wherever your data lives. As published on dataddo.com/products/data-anywhere.

4.1 ETL and ELT into your warehouse, lake or lakehouse

Most of the data an organisation needs for reporting and AI lives in SaaS applications: CRM, marketing, finance, product analytics. Data Anywhere™ loads it into Snowflake, BigQuery, Databricks, Redshift, Microsoft Fabric, Azure Synapse or open table formats (Iceberg, Delta Lake, Hudi), either transformed on the way in (ETL) or loaded first and transformed in the warehouse (ELT). The result is one source of truth for BI, data science and every AI tool that reads the warehouse.

Connectors are fully managed: Dataddo owns the data contract, so API changes, schema drift and deprecated endpoints are handled for you (Section 2.7). Beyond ETL and ELT, the same platform runs CDC, streaming, reverse ETL, batch file delivery and direct connections to AI (Section 2.2), with the same connectors, monitoring and access control.

4.2 Agentic data streaming

Agents that act on data need changes within a second, not after the next batch. Dataddo reads committed changes from database transaction logs, with no polling, triggers or extra indexes on the production system, and delivers deletes as events. A CDC supervisor restarts from the tracked log position if anything fails.

THE KEY DIFFERENCE - STREAMING SPEED (SQL SERVER CDC, INTERNAL BENCHMARK)

351 ms

median latency (p50), from
source commit to visible to agent

568 ms

90th percentile latency (p90)

35,000

change events per second,
sustained

Log-based CDC is supported for PostgreSQL, MySQL, SQL Server, Oracle, MongoDB, SAP HANA, Aurora, RDS, Cloud SQL, AlloyDB and CockroachDB, plus Kafka, Azure Event Hubs and Google Pub/Sub. Agents consume the changes three ways: by querying the warehouse or lake, by being pushed ordered events, or by reading directly from SmartCache over MCP, REST or Apache Arrow.

4.3 Private deployment for sensitive data

The data plane runs where the data must stay. Records stay inside your perimeter and credentials stay in your vault; only control signals travel to the Dataddo control plane.

Deployment	Where	Typical data
Public cloud	AWS, Azure, Google Cloud	Low to moderate sensitivity
EU sovereign cloud	Hetzner, OVHcloud, STACKIT	Regulated or residency-bound: PII, financial, health
On-premises	Kubernetes, OpenShift, bare metal	Highly sensitive or restricted, including air-gapped environments

Table 9. Data plane deployment options by data sensitivity.

4.4 Enterprise features

Area	What is included
Identity and access	SSO via SAML 2.0 or OIDC through your identity provider, role-based access control
Encryption	In transit and at rest, with bring-your-own-key via AWS KMS, Azure Key Vault or HSM
Network isolation	Fully isolated data plane in your private cloud or on premises; the control plane never sees the payload
PII protection	Detection, masking, tokenization or redaction at ingestion; PII fields can be excluded entirely
Resilience	Multi-region deployments, automated failover and built-in redundancy
Observability	Run histories, payload inspection, proactive anomaly monitoring and end-to-end lineage
Interoperability	Catalogs (Collibra, OpenMetadata, DataHub), orchestrators (Airflow, Dagster, Prefect), monitoring and SIEM (Datadog, Splunk, Sentinel)

Table 10. Enterprise capabilities of Data Anywhere™.

4.5 Plugging into OpenAI's Data agent and Claude Dashboards

The two newest AI analytics tools both expect your data to already be in a warehouse or lakehouse, which is exactly where Data Anywhere™ delivers it.

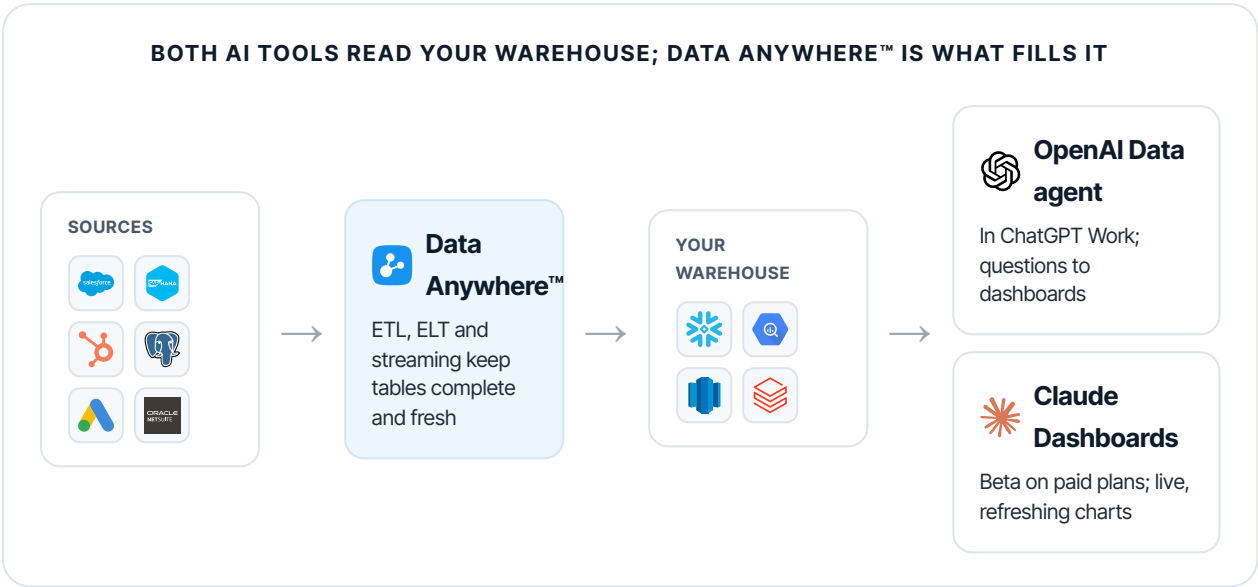


Figure 5. Data Anywhere™ as the feed for warehouse-native AI tools. Vendor announcements, September and October 2026.

	OpenAI Data agent	Claude Dashboards
Launched	September 10, 2026, in ChatGPT Work	October 8, 2026, beta on paid Claude plans
What it does	Turns plain-language questions into insights and dashboards across the company data stack	Builds a dashboard from a plain-language question and keeps it current; each chart shows its query and last refresh time
Data platforms	Amazon Redshift, Google BigQuery, Databricks, Snowflake	Amazon Redshift, BigQuery, ClickHouse, Databricks, Snowflake, plus connectors such as Salesforce
Data Anywhere™ adds	Loads 450+ sources into those platforms and keeps them fresh	The same, and streaming keeps "current as data changes" true to the second

Table 11. The two AI analytics tools and what Data Anywhere™ adds to each.

Both tools query data platforms plus a short list of app connectors. They do not consolidate hundreds of business applications and operational databases into one modeled, governed schema, so a dashboard is only as complete and current as the tables underneath it. The setup is therefore:

1. Connect your sources in Data Anywhere™ and choose your warehouse or lakehouse as the destination.
2. Use ETL or ELT for slowly changing data and agentic data streaming for tables that must be current to the second.
3. Grant the AI tool read access to the curated schema, following its own setup for Snowflake, BigQuery, Redshift or Databricks.
4. Optionally, add Dataddo's MCP server as a connector in Claude or ChatGPT, so assistants can also query SmartCache data and Dataddo's business metadata directly.

This keeps your warehouse as the single source of truth, under your own modeling and access rules, while every AI tool reads the same governed tables.

4.6 Commercial model and fit

Data Anywhere™ is priced per data flow, with enterprise terms for private deployments. It suits data and platform teams that already run Snowflake, BigQuery, Redshift, Databricks or a lake, need real-time feeds for agents, or must keep regulated data inside their own cloud or data center.

5 Decision guide

Most companies can decide with two questions.

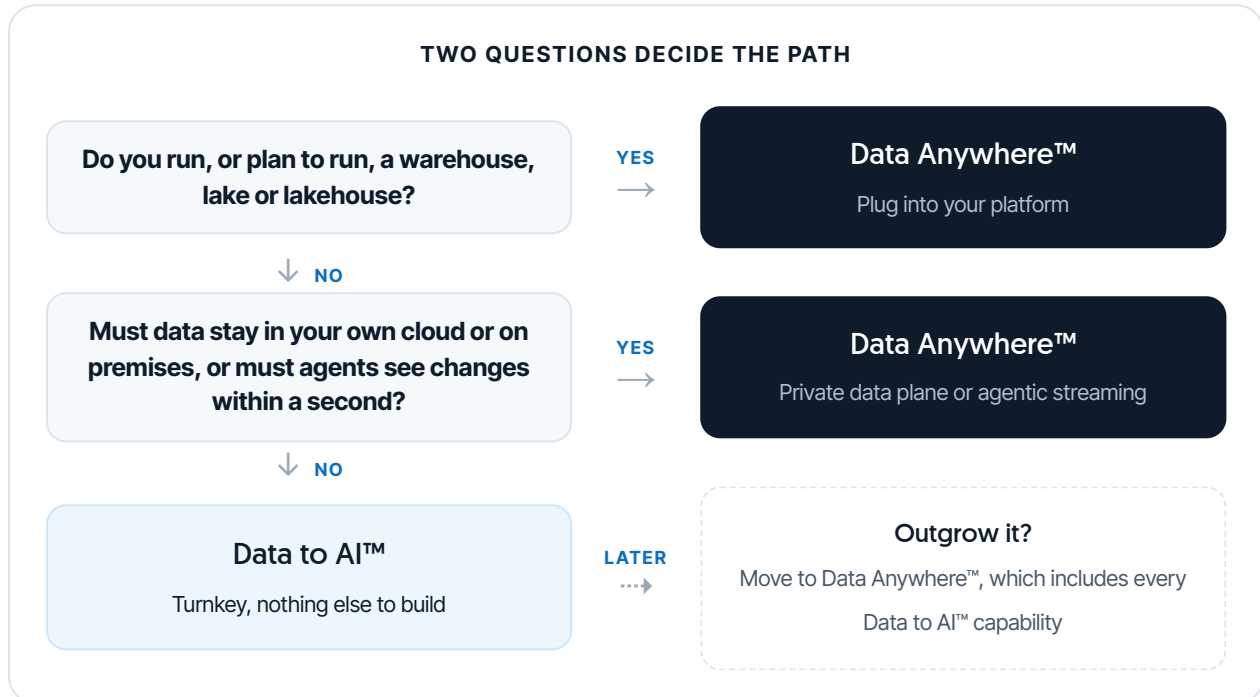


Figure 6. Decision guide: two questions, two outcomes, one upgrade path.

Three signals point the same way in practice:

- **Choose Data to AI™** when the people asking questions are business users, the data lives mainly in SaaS tools, and nobody wants to run a warehouse just to use AI.
- **Choose Data Anywhere™** when a data team already owns a platform and its rules: curated or staging layers, a semantic layer, an AI tool chosen centrally. A direct path from source to AI would bypass those rules.
- **Choose Data Anywhere™** when compliance requires a sovereign cloud, on-premises or air-gapped processing, or when agents must act on operational changes in under a second.

The choice is not one-way. Because Data Anywhere™ includes every Data to AI™ capability, a team can start with Data to AI™ for quick wins and move to Data Anywhere™ when it adds a warehouse, without changing AI vendors.

6 Side-by-side comparison

Data to AI™ buys a finished outcome; Data Anywhere™ buys a building block for an architecture you own.

Dimension	Data to AI™	Data Anywhere™
What you get	A complete service from source to AI answer	Connectivity and data movement into your own platform, plus every Data to AI™ capability
What else you need	Only your data sources and an AI client	A warehouse, lake or lakehouse, and a team to model and operate it
Data movement	Managed sync from 450+ sources into SmartCache	ETL, ELT, CDC, streaming, reverse ETL and file delivery from 450+ sources
Storage	Managed SmartCache, no warehouse	Your Snowflake, BigQuery, Redshift, Databricks, Fabric or object storage
Governance and lineage	Built in: quality firewall, business metadata, lineage, audit trail	Central policies and column-level lineage, integrating with Collibra, OpenMetadata or DataHub
Security	PII hashing and masking, row and column policies, read-only, encryption	All of that, plus SSO, RBAC, BYOK and private data planes
Deployment	Dataddo cloud	Dataddo cloud, EU sovereign cloud, your cloud or on premises, including air-gapped
Data freshness	Managed sync	Batch, or agentic data streaming at 351 ms median latency
AI clients	Claude, ChatGPT, Gemini, Copilot, Mistral, Grok, Cursor over MCP	The same over MCP, plus any tool that reads your warehouse, such as OpenAI Data agent and Claude Dashboards
Skills required	None beyond using an AI assistant	Data engineering for the platform; no code for the pipelines
Time to value	Authorize sources, add one connector in your AI client; no build phase	First pipelines live in hours; the overall timeline depends on your platform
Commercial model	Per seat; no platform to build or run	Per data flow; enterprise terms for private deployment

Table 12. Data to AI™ and Data Anywhere™ compared.

7 Appendix: methodology and sources

Benchmark	Method	Scope and limits
Answer quality (Data to AI™)	43 questions with ground truth written before any run, over Google Search Console, HubSpot deals and Google Ads data. The same model received the data four ways: CSV, CSV with data dictionary and SQL engine, generic SQL connector, Data to AI™ over MCP. Each question asked at least 3 times per condition; 759 graded runs.	Run by Dataddo on its own product; 7 questions deliberately unanswerable; models included Claude Sonnet 5, Claude Opus 5 and a low-cost open model.
Agentic data streaming latency	Time from commit at the source to the change being visible to the agent, on Dataddo's SQL Server CDC engine.	Internal benchmark; results for other engines and hardware may differ.

Table 13. Benchmarks cited in this paper.

This whitepaper was produced by Dataddo. Product capabilities and certifications are as published on dataddo.com in October 2026 and may change. Architecture schemas are reproduced from dataddo.com. Third-party product names and logos belong to their owners and are used to identify supported systems.

Put your own data behind your AI this week

Start with Data to AI™ on the sources your teams already use, or talk to us about plugging Data Anywhere™ into your warehouse, lake or lakehouse.

Try Data to AI™

Explore Data Anywhere™

Sources

All pages accessed October 11, 2026.

- [Dataddo Platform](#)
- [Dataddo Control Plane](#)
- [Dataddo for Claude](#)
- [Dataddo Data to AI™](#)
- [Dataddo Data Anywhere™](#)
- [Dataddo Agentic Data Streaming](#)
- [Dataddo, "Same data, same model. The data layer decides the answer." Benchmark report, September 2026](#)
- [Cloudera, Data Readiness Index 2026](#)
- [Model Context Protocol, What is MCP?](#)
- [Anthropic, Build live dashboards and animate explainers with Claude](#)
- [OpenAI Data agent launch report \(KuCoin News\)](#)