PRIVATE MARKETS

Databricks Investment Thesis

As companies move from experimentation to production, Databricks is positioning itself as the platform layer that unifies data governance, analytics and AI operations, turning fragmented information systems into a scalable foundation for enterprise intelligence.

KEY TAKEAWAYS

01	Databricks is positioned at the centre of enterprise AI adoption by unifying data, analytics and machine learning workflows.
02	Its revenue growth, 140%+ net revenue retention and expanding AI portfolio suggest a powerful consumption flywheel.
03	As AI moves from prototyping to production, platforms governing enterprise data could capture disproportionate value.

Every major wave of enterprise technology tends to create a new platform layer that becomes essential to enterprise operations. Mainframes had operating systems. The internet era had databases and middleware. The artificial intelligence (AI) era demands something different: a unified data and intelligence platform capable of governing, processing, and learning from an organisation's entire information estate. ARK believes Databricks is building that platform, and the economic evidence suggests that it is scaling faster than consensus appreciates.

The company's timing also aligns with a structural inflection in enterprise AI adoption. Analysts estimate that global IT spending will reach $5.43 trillion in 2025, with AI infrastructure demand driving much of that growth.[1] That said, spending on AI models alone does not generate enterprise value. Organisations must prepare, govern, and operationalise data before any model can produce reliable output.

Databricks addresses that structural bottleneck. Its "lakehouse" architecture collapses what historically were separate systems—data warehouses for analytics, data lakes for engineering, and bespoke environments for machine learning—into a single platform built on open data formats and consumption-based pricing. ARK's research indicates both that AI training costs have declined ~70% annually and that we are still in the early stages of AI use cases reaching production. As training and inference become less expensive, more enterprises can justify deploying AI workloads, driving greater consumption on the data platforms that prepare and govern the underlying data. As enterprises shift from AI prototyping to governed, production-grade deployment, the platform controlling data governance and model operationalisation should capture disproportionate value.

Databricks' revenue trajectory reflects that conviction, reporting $2.6 billion in fiscal year revenue ending January 2025 and expected annualised revenue to reach $3.7 billion by mid-2025—roughly 50% year-over-year growth.[2] By early 2026, what we believe is credible reporting placed the company's annual revenue run-rate above $5.4 billion, with year-over-year growth exceeding 65%.[3] For context, Snowflake reported $3.6 billion in fiscal year 2025 revenue, and net revenue retention (NRR) of 126%. Databricks has disclosed NRR exceeding 140% across multiple periods,[4] meaning existing customers expand consumption far more aggressively than its public peer. This is the hallmark of a consumption flywheel: once an organisation centralises data engineering workloads on Databricks, it naturally expands into SQL analytics, then to machine learning, then to generative AI, with each new workload consuming additional compute units within the same governed environment.

The composition of that expansion reveals Databricks' multi-vector growth strategy. Databricks' SQL business alone was targeting a $1 billion run-rate by fiscal year-end January 2026, up from $600 million a year prior.[5] Simultaneously, its AI product portfolio reached a $1 billion run-rate.[6] That dual-engine growth, displacing legacy data warehouses while attaching new AI workloads, creates nonlinear scaling dynamics. Databricks acquired MosaicML in 2023 as enterprise demand for custom LLM training accelerated, securing AI training and inference expertise that positions it to serve enterprises seeking tailored model development. Unity Catalog, the company's centralised governance system for data and AI assets, serves as a single surface for access control, auditing, lineage, and discovery across every workload.

The customer cohort data reinforces enterprise commitment. Databricks reported ~15,000 customers, with 800 spending more than $1 million annually and 70 exceeding $10 million.[7] Those are infrastructure-grade commitments. Moreover, its strategic acquisitions—MosaicML for large language model (LLM) training, Tabular for open table format leadership, and Neon for Postgres-based transactional capabilities—

signal Databricks' deliberate expansion from analytics into operational database territory, positioning the company to serve emerging agentic AI application patterns that require both analytical and transactional data access.[8] We believes that such execution velocity reflects the leadership of CEO Ali Ghodsi, Ph.D., and a founding team that created Apache Spark, Delta Lake, and MLflow open-source projects that anchor developer workflows across the industry.

The asymmetric case rests on what happens when a consumption-based platform with 140%+ NRR continues compounding across a customer base still early in its AI adoption curve. Enterprise AI workloads are becoming core operational infrastructure. ARK's research suggests that AI tools already improve knowledge-worker productivity by roughly two times in domains like AI-assisted coding, and we believe those productivity gains will expand as models improve and cost declines continue. The platform governing and processing the data beneath those workloads benefits from every incremental model deployment, every new agent, every additional governed dataset. Open-source ecosystem gravity around Delta Lake and MLflow anchors developer workflows in Databricks-originated standards, extending the company's relevance beyond any single product cycle.

Several structural risks warrant attention. Databricks remains private, limiting visibility into its audited margins and detailed unit economics. Microsoft Fabric's rapid adoption of 21,000 organisations within 18 months represents genuine bundling pressure.[9] Consumption models can decelerate during enterprise cost optimisation, and execution in the transactional database adjacency remains unproven at scale.

ARK's research indicates that Databricks represents a defining platform for the AI era, one whose economic flywheel is accelerating as enterprises move from AI experimentation to production-scale deployment. In our view, this convergence of data governance, analytical processing, and AI operationalisation on a single open platform represents a structural shift that the broader market has neither recognised nor priced fully.

References

1. TechRadar (citing Gartner). 2025. “Global AI Adoption to Push IT Spending beyond $5.4 Trillion in 2025.” TechRadar.

2. Novet, J. 2025. “Databricks Says Annualized Revenue to Reach $3.7 Billion by Next Month.” CNBC.

3. CRN. 2026. “Databricks Reports $5.4 Billion Revenue Run Rate as It Closes a $7B Investment Round.” CRN.

4. Novet, J. 2025. “Databricks Says Annualized Revenue to Reach $3.7 Billion by Next Month.” CNBC. CRN. 2026. “Databricks Reports $5.4 Billion Revenue Run Rate as It Closes a $7B Investment Round.” CRN. Lamba, K. et al. 2025. “Databricks Closes $1 Billion Round, Projects $4 Billion in Annualized Revenue on Surging AI Demand.” Reuters.

5. Ford, B. 2025. “Databricks Eyes $1 Billion in Sales for Product Competing with Snowflake.” Bloomberg.

6. Lamba, K. et al. 2025. “Databricks Closes $1 Billion Round, Projects $4 Billion in Annualized Revenue on Surging AI Demand.” Reuters.

7. CRN. 2026. “Databricks Reports $5.4 Billion Revenue Run Rate as It Closes a $7B Investment Round.” CRN. Lamba, K. et al. 2025. “Databricks Closes $1 Billion Round, Projects $4 Billion in Annualized Revenue on Surging AI Demand.” Reuters.

8. Lardinois, F. 2025. “Databricks to Buy Open Source Database Startup Neon for $1B.” TechCrunch. Novet, J. 2024. “Databricks Is Buying Data Optimization Startup Tabular.” CNBC.

9. Ford, B. 2025. “Microsoft Touts Sales in Competition with Snowflake, Databricks.” Bloomberg.

Important Information

This communication has been issued by IQ EQ Fund Management (Ireland) Limited ("**IQ EQ**") acting in its capacity as the manager of the funds referenced in this document. IQ EQ is authorised and regulated by the Central Bank of Ireland. IQ EQ is registered in Ireland with registration number 148223. If you have any questions regarding this marketing communication or in connection with the funds referenced in this document, you may contact IQ-EQ by email (ManCo@iqeq.com) or telephone (+353 1 673 5480). In the United Kingdom this communication is directed exclusively at, and intended only for, investment professionals under Article 19 of the Financial Promotions Order and high net worth companies, unincorporated associations and other persons falling within Article 49 of the Financial Promotions Order. This communication is not intended for retail investors. This communication is provided for information purposes only. The content of this communication constitutes general market, economic or thematic research and does not constitute investment advice, a personal recommendation, an offer, or a solicitation to buy or sell any financial instrument or investment product. The information presented does not take into account the individual investment objectives, financial situation, or particular needs of any specific person. Any opinions, forecasts and/or estimates expressed in the communication are sourced from ARK Invest International Ltd, ARK Investment Management LLC and their corporate affiliates ("**ARK**"), and such opinions, forecasts and/or estimates may be derived from third party information sources believed to be reliable. ARK does not make any representations or warranties, either express or implied, as to the accuracy or completeness of any such opinions or estimates or any information procured from any third-party sources. Any opinions, forecasts and/or estimates herein constitute a judgement that is subject to change without notice and do not represent a guarantee of future performance. ARK disclaims all liability and responsibility arising from any reliance placed by any person in the communication or on the opinions, forecasts and/or estimates and other information contained in this communication. The information contained in this communication is not intended to be complete or exhaustive and ARK does not make any representation or warranty, either express or implied, regarding the accuracy or completeness of the information. Reliance on the information is at the sole discretion of the reader and/or viewer. Investing involves risks. You should seek professional investment advice before making any decision to invest. Figures refer to the past and past performance is not a reliable indicator of future results. Where simulated past performance is shown, it must be noted that simulated past performance does not represent actual past performance and is not a reliable indicator of future performance. The value of investments can go down as well as up and investors may not get back the amount originally invested. Any forward-looking statements are based on assumptions and current expectations and may not prove to be accurate.