

Agentic Data Operations

Why Data Quality in Private Capital
is a Systemic Risk Problem—and How to Solve It

Steven Kilby, Chief Technology Officer
Jack Levy, Head of Data Strategy

Content Index

003 Executive Summary

004 The Systemic Risk Problem

006 NAV Lending: Where the Problem Is Most Acute

008 Why Agentic Data Operations Is the Answer

010 The Allvue Approach

012 What This Means for Private Market Practitioners

Executive Summary

Private capital markets manage over \$13 trillion in assets globally, yet operate with less data transparency than almost any other segment of institutional finance.

In February 2026, Bloomberg described these markets as “data black holes” that leave regulators unable to identify systemic risks.^[1] This is not an operational inconvenience. It is latent credit risk embedded in every portfolio, every facility, and every lending decision that depends on data the industry cannot yet reliably produce.

Three forces are converging to make this untenable. First, the explosive growth of private credit and fund-level leverage: NAV lending alone exceeds \$150 billion and continues to expand.^[2] Second, regulators across the US, UK, and Europe now expect banks to demonstrate consolidated visibility into private-market exposures.^[3] Third, AI systems are beginning to act autonomously on this data — pricing loans, monitoring covenants, generating risk scores. When AI agents make decisions based on bad data, the result is not a wrong number on a report; it is a machine-speed financial failure.

Against this backdrop, a new risk reality is becoming clear: scaling data quality through human labor inherently introduces fragility. Human processes are constrained by time, attention, turnover, inconsistency, and manual error. As volumes

grow and regulatory expectations tighten, these constraints compound — creating operational bottlenecks, inconsistent controls, delayed issue detection, and elevated compliance risk. Data quality managed primarily through headcount becomes reactive, episodic, and dependent on individual judgment rather than systematic enforcement.

Agentic Data Operations fundamentally reduces this risk profile. Instead of relying on manual intervention, AI agents continuously monitor, validate, enrich, and govern data in real time. They operate without fatigue, enforce policies consistently, adapt to drift automatically, and maintain comprehensive audit trails by design. At scale, data quality shifts from a people-dependent reconciliation exercise to an always-on control function — materially lowering operational risk while increasing transparency, resilience, and governance confidence.

This paper examines the data challenges facing private capital and explains why Allvue Systems is building its data products on a fundamentally different foundation: proprietary, rights-cleared data sourced directly from the operational systems of nearly 600 private capital firms; an AI-first operating model that delivers quality through software, not labor; and an architecture for engineering trust that makes every data point auditable, traceable, and ready for the AI-driven workflows that institutional investors increasingly demand.

AI without proprietary data is feature-level differentiation. AI built on top of proprietary data is platform-level differentiation. That distinction is the foundation of everything Allvue is building.

The Systemic Risk Problem

Private markets evolved without standardized disclosures, central exchanges, or uniform reporting requirements. Each general partner reports differently, on different timelines, using different definitions. Approximately 80% of enterprise data is unstructured, and the proportion in private markets is likely higher — valuable information trapped inside PDFs, emails, and spreadsheets.^[4]

For decades, this opacity was tolerable. Private capital was a niche allocation held by sophisticated investors comfortable with manual workarounds. That era has ended.

The Scale Has Changed

Private credit has moved from the periphery to the center of institutional portfolios. BlackRock's 2026 outlook described it as core to accessing growth and liquidity.^[5] McKinsey confirmed that allocations continue to increase even as operational infrastructure lags.^[6] Banks' total loan commitments to private equity and credit fund sponsors reached approximately \$300 billion by 2023, a thirty-fold increase from a decade earlier.^[7] Each of these exposures demands granular, continuously monitored data on thousands of underlying assets — data that largely does not exist in standardized form.

Regulators Are Paying Attention

The European Central Bank documented in 2025 that banks' risk management of private-market exposures tends to be fragmented — governed by product type, client segment, or business line — meaning that even the institutions lending to these markets often lack a consolidated view of their own exposure.^[8] The UK's Prudential Regulation Authority conducted a thematic review of private equity financing activities and sent a formal letter to bank chief risk officers.^[9] The Boston Federal Reserve published research quantifying the scale of bank lending to fund sponsors.^[10] The message from supervisors is consistent: the current state of data visibility is not acceptable.

The Stakes Are Higher Than They Appear

Poor data in private capital is not an operational nuisance. It is latent credit risk. When a lender cannot independently validate the collateral backing a NAV facility, that is not an administrative gap — it is a blind spot in a credit portfolio. When an insurer cannot benchmark covenant terms against market norms, that is not inefficiency — it is unquantified exposure.

And as AI agents begin pricing loans, generating risk signals, and executing covenant alerts, the consequences of bad data accelerate. Data quality failures become machine-speed financial failures. The tolerance for imprecise, stale, or inconsistent data shrinks toward zero.

NAV Lending: Where the Problem Is Most Acute

If the systemic risk problem in private capital is abstract, NAV lending makes it concrete. It is the clearest example of why better data infrastructure is not optional.

What Lenders Actually Face

A bank providing a NAV facility to a private credit fund is lending against a portfolio that may contain hundreds or thousands of individual loans. Each loan has its own borrower, its own terms, its own covenants, and its own performance trajectory. The lender must continuously monitor loan-to-value ratios, portfolio diversification limits, borrower concentration thresholds, and covenant compliance — all of which require aggregating data from multiple sources, often received in inconsistent formats.

A bank receiving loan tapes from ten different general partners will typically receive ten different data schemas, with inconsistent naming conventions, missing fields, and varying levels of granularity. Valuations of underlying assets must be obtained and validated on an ongoing basis. In stressed markets, there is scope for delays, disputes, and scenarios where lenders formally challenge the numbers they receive.

Scale Makes Manual Monitoring Impossible

NAV lending is not a niche product. Outstanding loans exceed \$150 billion, with the market having grown at approximately 30% annually between 2019 and 2023.

^[11] In 2025, a record \$12.9 billion in NAV lending funds closed, including the largest single fund in the asset class's history.^[12]

At this scale, spreadsheet-based monitoring introduces delays, inconsistencies, and blind spots that create real credit risk. Thousands of positions must be evaluated continuously. The regulatory environment reinforces the urgency: the Prudential Regulation Authority's 2024 thematic review, the ECB's dedicated monitoring exercise, and the Boston Fed's research all signal that supervisors expect better visibility.^[13]

Carlyle's 2026 credit outlook observed that many recent high-profile credit issues were driven not by market conditions but by weak documentation, poor diligence, or lack of lender control.^[14] The data gap is not an operational gap. It is a performance gap.

What Effective Monitoring Requires

The standard a NAV lender should expect: any loan in a collateral portfolio can be benchmarked against a relevant peer cohort — by sector, strategy, size, or structure — automatically and continuously.

Early warning signals are detected before covenant breaches occur. Loan tape inconsistencies are flagged and reconciled in near-real-time. Collateral quality, advance rates, and concentration metrics are assessed against anonymized market data rather than evaluated in isolation.

Achieving this requires three capabilities that do not currently exist at scale in private markets: standardized, rights-cleared, asset-level data that serves as a benchmark reference; automated quality checks that adapt to the inherent messiness of real-world loan tapes; and continuous monitoring that flags problems before they become breaches.

This is precisely the infrastructure Allvue has built.

Why Agentic Data Operations Is the Answer

The traditional approach to data quality — human experts writing validation rules, software enforcing them, human reviewers handling exceptions — was designed for a world where data arrived in batches and was consumed by human analysts. In private capital, that model has been stretched past its limits.

Data arrives continuously from hundreds of sources. Formats change without warning. Patterns drift gradually in ways that static rules cannot detect. And the cost structure is linear: more data requires more reviewers. Incumbent data providers in financial services often employ hundreds or thousands of manual reviewers — an approach that is expensive, slow, inherently inconsistent, and structurally unable to keep pace with the growth of private markets.

A Fundamentally Different Model

Agentic data operations replaces static rules with adaptive systems. AI agents learn what normal data looks like by analyzing historical patterns — volume, distribution, freshness, completeness, relationships between fields. They detect anomalies and drift autonomously, without requiring pre-programmed thresholds. They propose or execute corrections. They maintain full audit trails. And humans shift from performing quality checks to governing quality systems.^[15]

In concrete terms, agentic data operations is the mechanism by which:

Opacity becomes measurable. Data that was previously unstructured and inaccessible is standardized, quality-scored, and made available for analysis.

Quality becomes scalable. Instead of adding reviewers to match data growth, the system improves with every dataset it processes. Each new record strengthens the models that govern all records.

Governance becomes continuous. Rather than periodic manual audits, AI agents enforce data policies in real time — tracking lineage, managing access rights, ensuring de-identification protocols, and documenting every transformation.^[16] The Savant 2026 Trends Report, which surveyed senior Finance, Tax, and Accounting leaders across 22 industries, reinforces this point: data governance and security was cited as the single largest obstacle to scaling AI by 37% of respondents — outranking cost, skills gaps, and workforce resistance combined. Savant's conclusion mirrors the principle embedded in Allvue's architecture: governance is not a downstream implementation detail. It is the primary gate to adoption.^[17]

Risk becomes observable in real time. Anomaly detection, drift monitoring, and cross-source consistency checks replace manual spot checks. Problems are caught before they reach downstream consumers — whether that consumer is a human analyst, a compliance report, or an AI model making credit decisions.

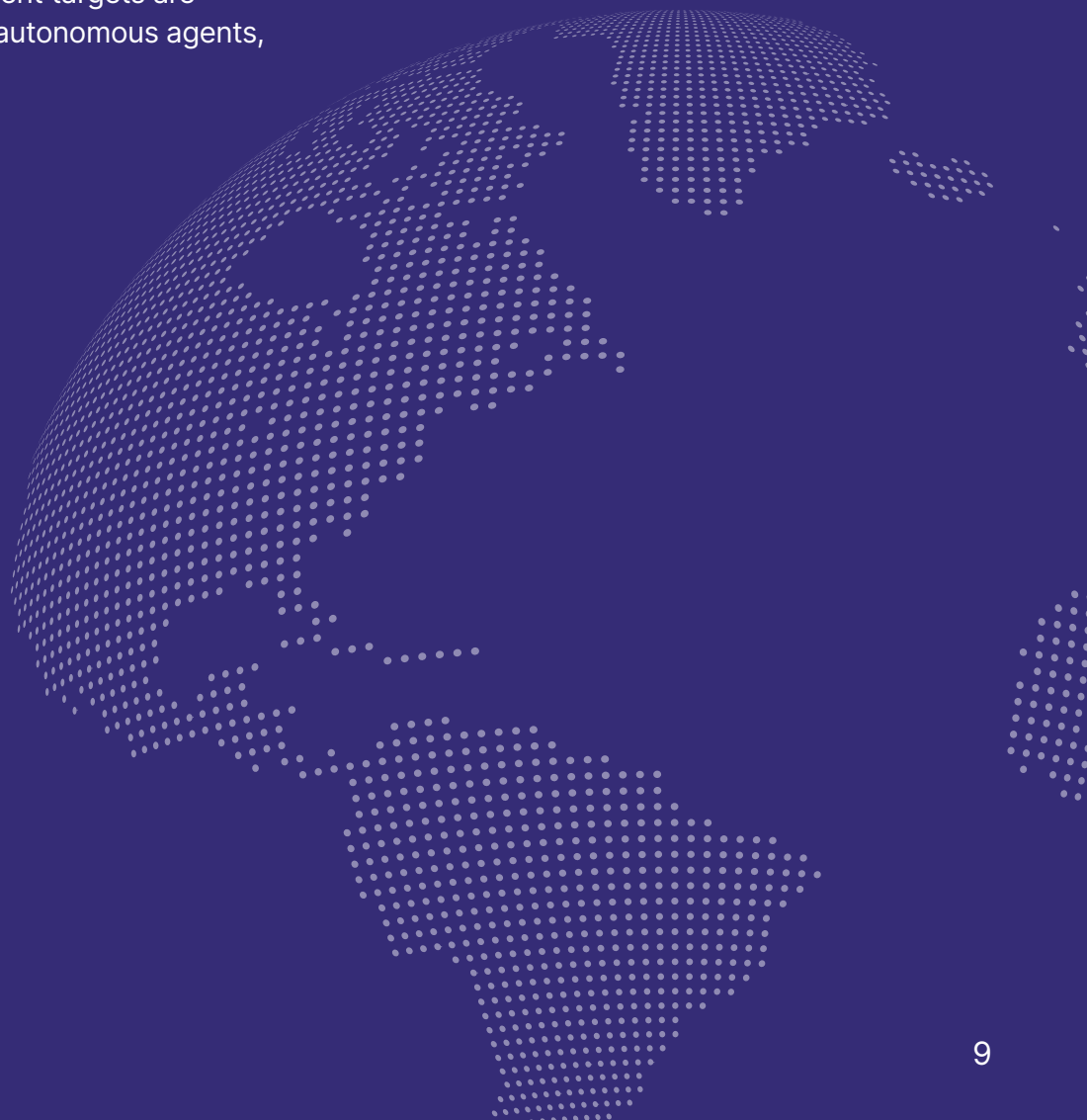
The Industry Recognizes the Shift

In February 2026, Gartner rebranded its flagship evaluation from the Magic Quadrant for Data Quality Solutions to the Magic Quadrant for Augmented Data Quality Solutions — a signal that AI-driven quality is now the expected baseline.^[18] IBM's positioning around its data intelligence platform emphasizes moving from reactive remediation to continuous, proactive management of data trust.^[19] KPMG's Q4 2025 survey found that data quality concerns among executives rose from 37% to 65% as agentic AI deployments expanded, with 75% citing security, compliance, and auditability as the most critical requirements for agent deployment.^[20]

Savant's 2026 research confirms this trajectory among finance practitioners. When leaders were asked to rank expected benefits of AI and automation, 43% placed speed and efficiency first — not cost reduction (12%) or headcount savings. The primary investment targets are generative AI assistants and autonomous agents,

which together account for 76% of all planned spending. The message from finance leadership is clear: the goal is not cheaper operations but faster, higher-quality decisions made at scale.^[21]

The academic literature reflects the same trajectory. A January 2026 survey documented the shift from rule-based data preparation to prompt-driven, context-aware, and agentic preparation workflows.^[22] Gartner estimates that by 2028, 33% of enterprise applications will include agentic AI, up from less than 1% in 2024.^[23]



The Allvue Approach

Allvue provides enterprise software and data infrastructure to nearly 600 private capital firms, representing approximately \$9 trillion in assets under management or administration across more than 22,000 funds. That operational footprint is not incidental to the data business. It is the foundation of it.

Allvue's data products are built on three pillars that, in combination, are extremely difficult to replicate: proprietary rights-cleared data, AI-first data operations, and an architecture for engineering trust.

Proprietary, Rights-Cleared Data

The quality of any data product ultimately rests on the provenance of the underlying data. In private capital, this means data sourced directly from the systems of record where transactions actually occur — fund accounting platforms, portfolio monitoring tools, investor reporting workflows — rather than scraped from public filings, assembled from surveys, or inferred from secondary sources.

Allvue's data is drawn from the operational systems of its client base. The company holds contractual rights to approximately 70% of this dataset, with regular updates governed for quality, granularity, and compliance. De-identification, aggregation, and anonymization are managed through documented client permissions.

In private credit, Allvue maintains anonymized, rights-cleared data across more than 200 data points covering borrowers, facilities, transactions, and covenant terms. The dataset spans tens of thousands of loans

and borrowers, with coverage expanding substantially through the first half of 2026. The data is structured, mastered, and continuously refreshed for integration into credit, risk, and valuation workflows.

Market participants consistently tell data providers that they value systems-of-record data over survey-based or inferred alternatives — particularly for credit risk, valuations, and benchmarking, where accuracy is non-negotiable. AI without proprietary data is feature-level differentiation. AI built on proprietary, workflow-derived data is platform-level differentiation.

AI-First Data Operations

Allvue's Data Excellence team was designed from the start as an AI-first operating foundation for the Nexius data platform. The team is intentionally small — a handful of specialists whose mandate is not to manually review records but to build reusable systems that run continuously and scale with the dataset.

This design reflects where the broader market is heading. According to the Savant 2026 Trends Report, 82% of Finance, Tax, and Accounting leaders expect no net change in headcount driven by AI in 2026. The shift is not from people to machines. It is from linear scaling — where more data requires more reviewers — to compounding returns, where better systems make every person more effective. That is the operating philosophy behind Allvue's Data Excellence team.[24]

The approach includes automated validation and reconciliation logic that catches inconsistencies across loan tapes, general

partners, and reporting periods. Anomaly detection models flag outliers and sudden changes in reported values. Drift monitoring identifies slow degradation over time — the kind of creeping data quality erosion that static rules miss entirely. And controlled, auditable techniques close data gaps where reported values are incomplete, with every estimated value explicitly flagged and traceable.

Consider a practical illustration. A bank's NAV facility is backed by a diversified private credit portfolio. Each quarter, the general partner submits a loan tape. The system ingests that tape, normalizes the fields to a standardized schema, and runs it against the borrower's historical profile and its peer cohort — comparable borrowers in the same sector, strategy, and size band. If a borrower's reported leverage ratio has drifted meaningfully from prior quarters, or is inconsistent with the financial metrics of similar borrowers in the dataset, the system flags it automatically. The lender sees the anomaly before it affects a covenant calculation or a borrowing base determination — not weeks later in a quarterly review, but in near-real-time.

Each quality rule, enrichment algorithm, and governance control strengthens the dataset not just for a single use case, but for every current and future client. The breadth and depth of coverage expand materially — scaling in multiples rather than marginal increments — without compounding operational risk through human labor constraints.

Enterprises lose an estimated \$12.9 million annually due to poor data quality, and most of these failures are silent: systems rarely crash; they gradually degrade.^[25] Allvue's infrastructure is purpose-built to detect and remediate those degradations before they propagate — transforming data quality from a reactive clean-up effort into a proactive, systemic control function.

Engineering Trust

In private capital data, trust is the product. Every downstream consumer — whether a bank underwriting a NAV facility, a ratings agency calibrating a credit model, an insurer assessing portfolio risk, or an AI system generating a risk narrative — needs to know exactly where the data came from, how it was transformed, and whether it was directly reported or derived.

Allvue's governed semantic layer provides that foundation: a standardized schema and business glossary that ensures consistency across every dataset and use case; full metadata and lineage tracking that records every transformation a data point undergoes; and role-based access controls that restrict outputs to authorized, de-identified data.

The data becomes more reliable over time. Each new dataset processed by the system trains the anomaly detection models, refines the validation logic, and deepens the peer cohorts against which future records are evaluated. This compounding effect means that the quality of Allvue's data products improves continuously — not because more people are reviewing records, but because the infrastructure itself is learning.

Emerging across the industry is the concept of a data trust score: a composite, quantitative measure of dataset reliability across freshness, accuracy, completeness, lineage integrity, and governance compliance. Rather than presenting data quality as a binary pass/fail, trust scores enable consumers — including AI systems — to make calibrated decisions about how much confidence to place in each data point. This is particularly important when AI agents are acting on data autonomously: they need to know not just the value of a metric, but how much to trust it.

What This Means For Private Market Practitioners

The private capital data landscape is changing rapidly. The institutions that will navigate this transition most effectively are those that recognize what has shifted — and choose their data partners accordingly.

If you are a NAV lender or fund finance provider: the scale and complexity of your monitoring obligations have outpaced manual processes. You need asset-level benchmarks, continuous anomaly detection, and independent covenant visibility — not quarterly spreadsheets. The data backing your facilities should be sourced from operational systems, not assembled from secondary sources. And it should be validated by the same AI-driven quality infrastructure that institutional-grade buyers demand.

If you are buying or licensing private capital data: provenance matters more than volume. Ask where the data comes from. Ask whether it is contractually rights-cleared. Ask how quality is maintained as the dataset grows. Ask whether derived or estimated values are explicitly flagged. The difference between data you can build on and data that introduces risk is governance, not size.

If you are deploying AI on private capital data: the foundation matters more than the model. AI systems trained on poorly governed, inconsistently sourced data will produce poorly governed, inconsistently sourced outputs — at scale and at speed. The organizations that build effective AI in private markets will be those whose data infrastructure is designed for auditability,

lineage, and trust from the ground up.

The Savant 2026 Trends Report offers a useful framing for the cost of delay: across 22 industries, organizations that remain in pilot mode risk structural disadvantage, while those that align strategy, governance, and execution pull ahead.^[26] In private capital, where data opacity already creates latent risk, the penalty for hesitation is steeper still.

Allvue's position at the center of private capital workflows — the fund accounting platforms, portfolio monitoring tools, and investor reporting systems where transactions actually occur — is what makes its data products fundamentally different from alternatives assembled from public filings, surveys, or inferred sources. The proprietary data, the AI-first quality infrastructure, and the architecture for engineering trust are not separate initiatives. They are a single, integrated system designed to deliver the institutional-grade data that private capital increasingly requires.

The question facing the industry is no longer whether AI will transform data operations in private markets. It is which data providers have the proprietary foundations, the domain expertise, and the institutional trust to lead that transformation.

For Allvue, the answer is clear. The infrastructure is built. The data is proprietary. The quality compounds. And the standard of trust is validated by the most demanding buyers in the market.

Sources

All sources cited in this paper are referenced by number throughout. A consolidated list is provided below in the order they appear.

1. Bloomberg, "Private Markets' Data Black Holes Leave Watchdogs Flying Blind," February 2026.
2. S&P Global / Citco Capital Partners, NAV lending market data, 2025.
3. European Central Bank Banking Supervision, "Hidden Leverage and Blind Spots: Addressing Banks' Exposures to Private Market Funds," June 2025.
4. BNY / Funds Europe, "How to Ensure Proper Data Management in Private Markets," January 2026.
5. BlackRock, "2026 Private Markets Outlook."
6. McKinsey & Company, "Global Private Markets Report 2025," May 2025.
7. Boston Federal Reserve, "Growth of Bank Lending to Private Equity and Credit Funds," 2025.
8. European Central Bank Banking Supervision, "Hidden Leverage and Blind Spots: Addressing Banks' Exposures to Private Market Funds," June 2025.
9. UK Prudential Regulation Authority, "Thematic Review of Private Equity Related Financing Activities," April 2024.
10. Boston Federal Reserve, "Growth of Bank Lending to Private Equity and Credit Funds," 2025.
11. S&P Global / Citco Capital Partners, NAV lending market data, 2025.
12. With Intelligence, "Private Credit Outlook 2026," January 2026.
13. UK Prudential Regulation Authority, "Thematic Review of Private Equity Related Financing Activities," April 2024.
14. Carlyle, "Credit in 2026: A Market That Demands Insight, Not Just Capital."
15. Databricks, "Data Quality Monitoring at Scale with Agentic AI."
16. VentureBeat, "The Era of Agentic AI Demands a Data Constitution, Not Better Prompts," February 2026.
17. Savant Labs, "2026 Trends Report: Agentic Automation in Finance, Tax, and Accounting," February 2026.
18. Gartner, Magic Quadrant for Augmented Data Quality Solutions, February 2026.
19. IBM, "IBM Named a Leader in the 2026 Gartner Magic Quadrant for Augmented Data Quality Solutions," February 2026.
20. KPMG Q4 2025 AI Pulse Survey: 75% of respondents cited security, compliance, and auditability as the most critical requirements for agent deployment.
21. Savant Labs, "2026 Trends Report: Agentic Automation in Finance, Tax, and Accounting," February 2026.
22. Zhou et al., "Can LLMs Clean Up Your Mess? A Survey of Application-Ready Data Preparation with LLMs," arXiv, January 2026.
23. Gartner (via Acceldata): 33% of enterprise applications to include agentic AI by 2028, up from less than 1% in 2024.
24. Savant Labs, "2026 Trends Report: Agentic Automation in Finance, Tax, and Accounting," February 2026.
25. AnalyticsWeek, "Self-Healing Data Pipelines: Why 2026 Ends the Data Fire Drill," January 2026.
26. Savant Labs, "2026 Trends Report: Agentic Automation in Finance, Tax, and Accounting," February 2026.

ALLVUE 



**AI-powered
alternative
investment
software for
the full fund
lifecycle.**

Contact Us
www.allvuesystems.com