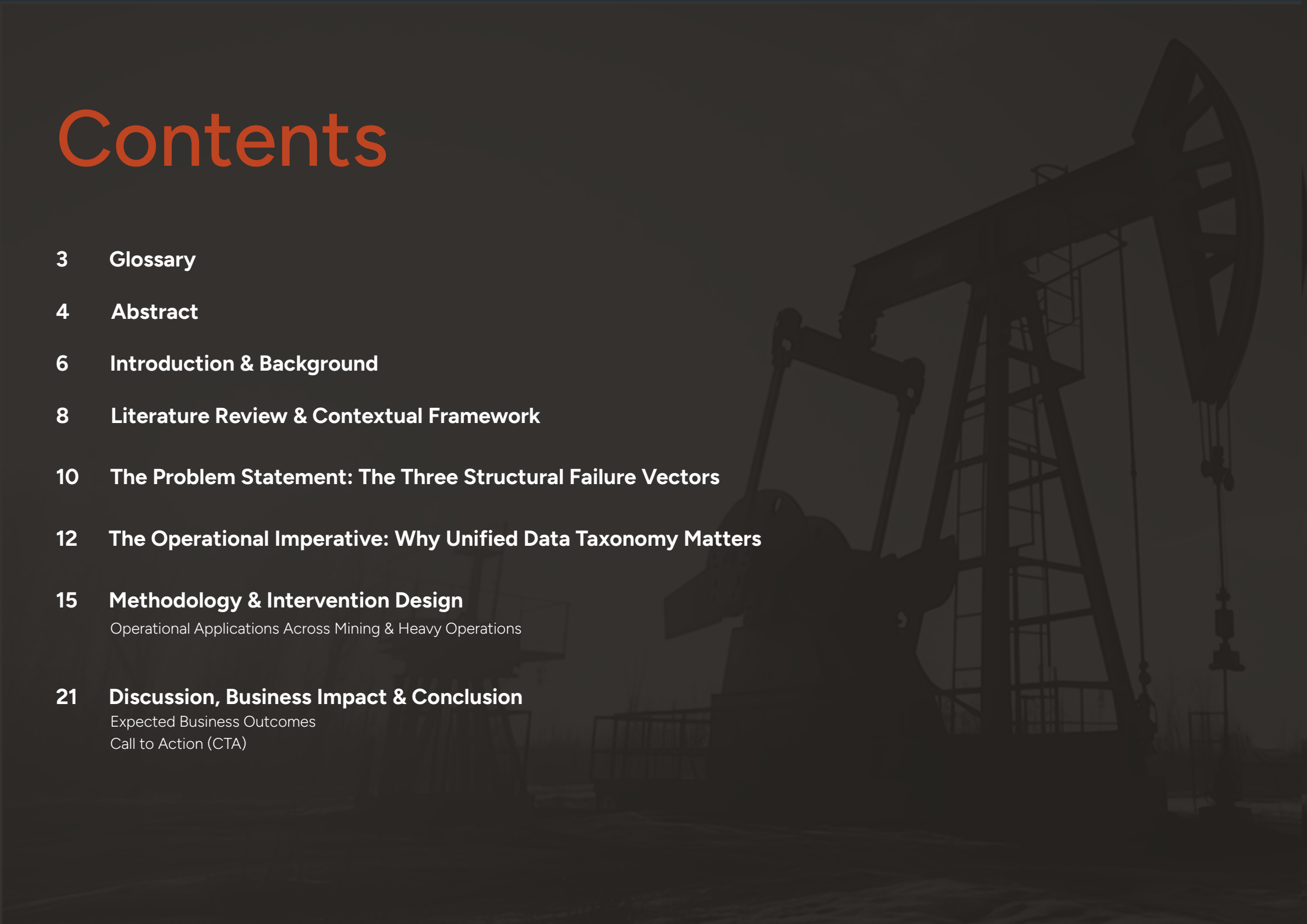


Fusing Edge Telemetry and Prescriptive AI to Drive Mining Operations



Contents



- 3 Glossary**
- 4 Abstract**
- 6 Introduction & Background**
- 8 Literature Review & Contextual Framework**
- 10 The Problem Statement: The Three Structural Failure Vectors**
- 12 The Operational Imperative: Why Unified Data Taxonomy Matters**
- 15 Methodology & Intervention Design**
 - Operational Applications Across Mining & Heavy Operations
- 21 Discussion, Business Impact & Conclusion**
 - Expected Business Outcomes
 - Call to Action (CTA)

Glossary

The Gen AI Paradox: Describes the structural disconnect where generative AI has achieved widespread enterprise adoption (~80%), yet yields minimal bottom-line earnings because deployments remain focused on diffuse horizontal tools (like chatbots and copilots) while high-impact vertical operational processes stay trapped in pilot mode.

F1-Score: A statistical metric measuring model accuracy as the harmonic mean of precision and recall. In technical log processing, it evaluates how accurately an NLP model extracts key operational entities, such as failure symptoms and machine actions, without false positives or omissions.

Case-Based Reasoning (CBR): An artificial intelligence framework that solves new diagnostic problems by finding analogous past events. Rather than relying solely on first-principle physics models, CBR matches incoming failure symptoms with historical maintenance logs to retrieve and adapt proven corrective actions.

Bidirectional Encoder Representations from Transformers (BERT): A transformer-based language model designed to understand the context of words in text by analysing surrounding words in both directions simultaneously, rather than sequentially.



Abstract

Abstract

Today, over 78% of enterprises have adopted artificial intelligence, more than 80% report no bottom-line earnings contribution, and fewer than 10% of vertical use cases scale beyond pilot phases. In capital-intensive mining operations, this Gen AI Paradox stems from deploying passive, horizontal copilots that deliver modest 5% to 10% task gains rather than orchestrating end-to-end execution. Critical operational intelligence remains trapped across three disconnected tiers: Tier 1 - Kinematic dispatch dynamics, Tier 2 - Physical machine telemetry, and Tier 3 - Contextual qualitative logs.

The Architectural Solution

To close this execution gap, our point of view outlines a closed-loop 5-Stage Data Architecture:



Edge Ingestion: On-site hardware captures live operational data directly from heavy machinery, including vehicle motion, engine diagnostics, component vibration, and GPS location.



Fusion & Translation: The system standardises conflicting equipment data and combines handwritten shift notes with live machine metrics into a single, unified operational timeline.



Semantic Inference: Specialised language models analyse technician failure notes and match them against historical breakdown records, raising diagnostic accuracy from 58% to 95%.



Prescriptive Closed-Loop Execution: Autonomous AI agents immediately create maintenance work orders, reserve replacement parts, and reroute haul trucks to prevent bottlenecks at processing plants.



Unified Governance: Built-in safeguards enforce risk policies, maintain safety compliance, and provide a clear audit trail for every automated decision.

Bottom-Line Impact

Transitioning from reactive alarms to autonomous agentic workflows compresses incident resolution cycle times by 60% to 90%, automates up to 80% of routine frontline triage events, and eliminates haulage bottlenecks that otherwise cause 12% to 25% productivity losses. Mining leadership must modernise enterprise architectures to orchestrate machine intelligence directly alongside human operations.

A miner wearing a hard hat and work clothes stands in a dark, rocky mine. In the center, a glowing orange digital interface with various charts and a central circular graphic is superimposed over the scene. To the right, a large pile of dark, jagged rocks sits in a metal container. The overall atmosphere is industrial and futuristic.

Introduction & Background

Introduction & Background

Operational Environment & Industry Context

Modern open-cast and underground extraction environments represent highly complex, capital-intensive systems where margins hinge on fleet availability, cycle-time efficiency, and zero-harm compliance. In these asset-heavy environments, operational resilience is no longer defined by sensor density, but by the capacity to orchestrate closed-loop execution directly around real-time field telemetry.

Yet, heavy industry remains constrained by the Gen AI Paradox: while over **78%** of enterprises have deployed artificial intelligence, more than **80%** report no material bottom-line contribution to earnings, fewer than 10% of high-impact vertical use cases scale beyond pilot phases, and only **1%** consider their AI operating model mature. Escaping this pilot purgatory demands moving past passive dashboards and embedding autonomous agentic systems directly into core extraction workflows, transforming high-velocity kinematic telemetry, fixed-plant vibrations, and geotechnical sensor feeds into immediate, synchronised interventions.

Despite extensive capital investment in edge monitoring, vast reserves of field intelligence remain what we call operational dark data. High-frequency machine signals stay locked inside proprietary controllers, while frontline qualitative logs remain trapped on paper check-sheets and unstructured shift handovers. Bridging this operational disconnect, fusing unstructured human observations with high-speed sensor streams into governed enterprise action, marks the critical frontier of industrial digital maturity.



Literature Review & Contextual Framework

Literature Review & Contextual Framework

Physical AI, Edge Computing, and Multi-Sensor Fusion

Physical degradation in heavy industrial machinery exhibits non-linear dynamics driven by thermal stress, shock loads, and mechanical fatigue. In distributed, high-noise environments, single-sensor monitoring suffers from blind spots and transient interference, requiring multi-sensor fusion at the edge:

High-Frequency Anomaly Detection: Fusing tri-axial vibration, current signatures, and acoustic emissions at the edge accelerates feature extraction, predicting Remaining Useful Life (RUL) before micro-faults trigger catastrophic failure.

Multimodal Edge Perception: Combining complementary physical sensor streams filters operational noise, contextualising raw telemetry for autonomous decision systems.

Kinematic Fleet Synchronisation: Discrete-event modeling shows that decoupled telematics in open-pit haulage induce spatial queuing and cycle-time variance, dragging productivity 12% to 25% below theoretical limits.

Semantic Abstraction & Unstructured Data Processing

While physical sensor streams capture mechanical wear, a vast repository of operational intelligence remains locked in unstructured text, including shift handover logs, operator field comments, and technician work orders. Standard Prognostics and Health Management (PHM) frameworks typically exclude this qualitative intelligence due to the semantic complexity of free-form industrial text.

Technical Language Processing (TLP): Standard NLP fails on fragmented industrial jargon. Adapting transformer models to maintenance syntax bridges the semantic gap across frontline narratives.

Transformer Token Classification: Domain-fine-tuned transformers (such as BERT) extract operational entities from free-form logs, yielding an 88% F1-score across failure symptoms, operational actions, and ambient conditions.

Textual Case-Based Reasoning (TCBR): Domain-fine-tuning Bidirectional Encoder Representations from Transformers (BERT) models on maintenance text via unsupervised sequential denoising allows historical logs to be converted into dense semantic embeddings. By pairing these embeddings with Case-Based Reasoning (CBR), systems match new failure descriptions against historical cases using cosine similarity, boosting diagnostic precision from 58% to 95% and enabling automated retrieval of proven repair actions.

Problem Statement



The Problem Statement: The Three Structural Failure Vectors

With fewer than 30% of enterprise AI agendas directly sponsored at the executive level^{2e}, field-level execution consistently breaks down across three core architectural vectors:

Operational Data Isolation: Proprietary telemetry locked inside mobile fleets and fixed plants creates spatial blind spots, causing haulage productivity, which represents 50% to 60% of operating costs, to slide 12% to 25% below theoretical limits.

The Unstructured Data Barrier: Disjointed edge hardware and free-text technician logs report identical mechanical faults through conflicting syntaxes, requiring domain-specific NLP models (achieving an 88% entity-recognition F1-score) to unlock actionable diagnostic data.

Alarm Fatigue & Maintenance Paralysis: In the absence of autonomous triage, control rooms are overwhelmed by unranked, context-free alarms that force frontline crews into reactive maintenance loops. Relying on superficial copilots yields only 5% to 10% efficiency gains, leaving critical incident resolution cycles extended by up to 90% longer than necessary under manual intervention.





The Operational Imperative: Why Unified Data Taxonomy Matters

The Operational Imperative: Why Unified Data Taxonomy Matters

In modern mining, operational disruptions rarely originate in isolation; rather, they cascade across kinematic, physical, and human operational layers. For example, excessive truck queuing at a loading bench (Tier 1 kinematic friction) often traces back to unpredicted crusher jaw wear (Tier 2 mechanical strain), which an operator had already noted days prior as abnormal grinding noise in a shift handover comment (Tier 3 unstructured log).

When these three layers remain siloed:

- SCADA systems detect physical stress too late to prevent catastrophic downtime.
- Fleet Management Systems (FMS) route haulers into congested zones without knowledge of downstream processing delays.
- Enterprise Asset Management (EAM) platforms miss the qualitative frontline observations that explain why equipment is failing.

Achieving autonomous, condition-based execution requires unifying these three distinct telemetry layers into a single, synchronised state space. Table 1 outlines this structural taxonomy, detailing how fragmented edge signals directly translate into operational and financial losses across the mining lifecycle.

The 3-Tier Traditional Mining Data Taxonomy

Tier 01

Architectural Data Tier

Kinematic & Spatial Dynamics
(Macro Movement & Topography)

Core Measured Parameters & Signals

Payload weight, vehicle velocity, fuel consumption rate, tire pressure/temperature, GPS coordinates, bench elevation, haul road gradient.

Primary Native Ingestion Sources

Fleet Management Systems (FMS), On-board GPS receivers, Mine GIS databases.

Structural Data Friction & Operational Consequence

High-frequency vehicle telematics remain decoupled from static pit topography and live traffic congestion, causing haul bottlenecks, queuing, and excessive fuel burn.

Tier 02

Architectural Data Tier

Physical & Mechanical Telemetry
(Micro Machine Physics)

Core Measured Parameters & Signals

Tri-axial vibration frequencies, hydraulic oil pressure, engine temperature, fluid contamination levels, crusher jaw pressure, conveyer belt speed.

Primary Native Ingestion Sources

CAN-bus controllers, SCADA systems, Programmable Logic Controllers (PLCs).

Structural Data Friction & Operational Consequence

Proprietary OEM protocols generate raw, high-volume alarm floods lacking failure context, forcing reliance on inefficient calendar-based maintenance.

Tier 03

Architectural Data Tier

Contextual & Environmental Signals
(Human & Ecological Layer)

Core Measured Parameters & Signals

Shift handover notes, pre-trip inspection sheets, technician voice memos, ambient dust density, blast vibration logs, slope stability radar

Primary Native Ingestion Sources

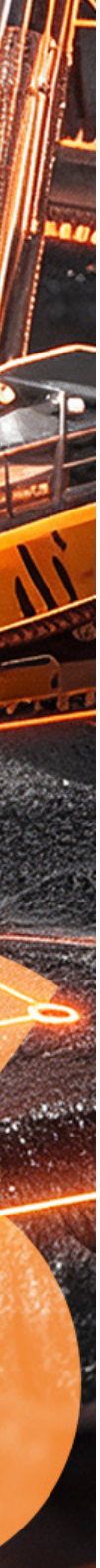
PDF work orders, mobile audio notes, paper check-sheets, Geotechnical radar, IoT weather stations.

Structural Data Friction & Operational Consequence

Critical qualitative observations and safety telemetry remain trapped in unstructured text or isolated silos, unread by automated analytical engines



Methodology & Intervention Design



Methodology & Intervention Design

The intervention deploys a 5-Stage Data Architecture Framework to convert raw physical edge signals and qualitative field intelligence into governed, autonomous enterprise actions:

Stage 1: Ingestion Layer (Edge IoT & Physical Telemetry Capture): Establishes a hardware abstraction interface across stationary plant infrastructure and mobile mining fleets using secure industrial gateways (e.g., Cisco industrial routers, Azure IoT Edge runtimes). It continuously captures Tier 1 Kinematic telematics (payload mass, truck velocity, tire temperatures, high-precision GNSS coordinates) and Tier 2 Physical streams (high-frequency CAN-bus data, tri-axial vibration spectra, hydraulic pressure, and passive/active RFID inventory logs) directly at the machine interface. To scale this layer, the technical engineering team will have to enhance edge hardware bandwidth optimisation, driver abstraction protocols, and secure firmware lifecycle management.

Data ingestion is managed through containerised AWS IoT Greengrass runtimes deployed on industrial edge gateways (e.g., Cisco IR1101). High-velocity CAN-bus signals (J1939 protocols) and tri-axial vibration metrics are captured via MQTT/HTTPS endpoints, while unstructured shift logs and voice recordings are staged in Amazon S3 Express One Zone for low-latency retrieval.

Stage 2: Fusion & Translation Layer (Sensor Fusion & Semantic Abstraction): Edge compute nodes normalise disparate OEM fieldbus protocols and parse unstructured Tier 3 contextual text, including shift logs, operator notes, and technician work orders^{4g}. By applying Technical Language Processing (TLP) alongside multimodal sensor fusion, the system aligns discrete machine physics, pit topography, and qualitative operator observations into a unified, timestamp-synchronised JSON payload. To ensure seamless multi-source alignment, the technical engineering team will have to enhance domain tokenisation pipelines, latency-bounded stream synchronisation, and edge serialisation routines.

Stream processing is handled by Amazon Kinesis Data Streams paired with AWS Lambda execution workers. Disparate OEM fields and unstructured free-text logs are normalised into a unified schema using custom Technical Language Processing (TLP) tokenisers. The parsed payload is structured into timestamp-synchronised JSON streams and persisted in an Amazon DynamoDB operational data lake cache for real-time state management.

Stage 3: Inference Layer (Predictive Machine Learning & Case-Based Reasoning): Lightweight ML models deployed at edge nodes alongside centralised cloud clusters evaluate normalised operational feature vectors to compute dynamic Remaining Useful Life (RUL) on critical mechanical components. Simultaneously, fine-tuned BERT models generate dense semantic embeddings of extracted failure descriptions, feeding a TCBR engine that matches active failure signatures against historical resolution records using cosine similarity at up to 95% precision. To maintain high diagnostic accuracy, the technical engineering team will have to enhance continuous model drift monitoring, unsupervised transformer re-tuning workflows, and high-dimensional vector search indexing.

Machine learning inference runs on Amazon SageMaker endpoints hosting RUL predictive models. Simultaneously, domain-adapted BERT transformer models generate dense semantic embeddings from unstructured failure text. These embeddings feed an Amazon Bedrock / OpenSearch Vector Engine, executing TCBR via cosine similarity matching at up to 95% diagnostic precision.

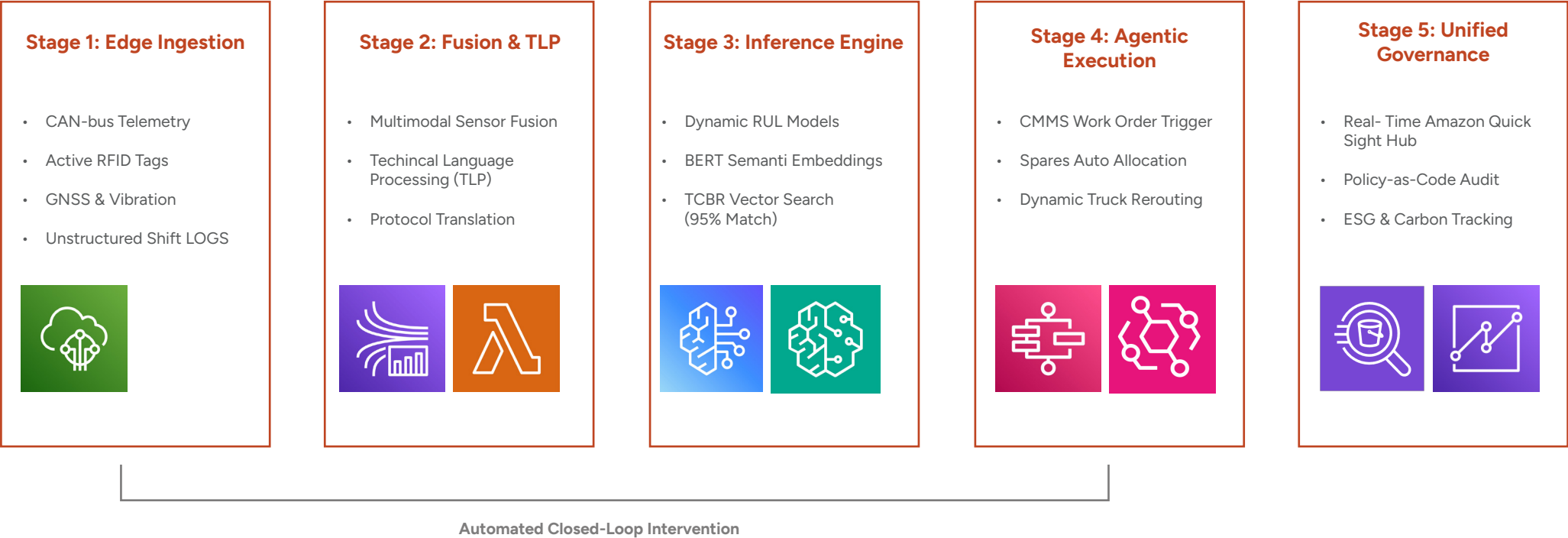
Stage 4: Execution Layer (Prescriptive Agents & Closed-Loop Workflows): Predictive inference scores and retrieved resolution cases are translated into deterministic operational actions via autonomous agentic orchestrators and low-code integration runtimes (e.g., Microsoft Power Platform, Power Automate, Copilot Studio). Rather than emitting passive alarm floods to congested control room consoles, the agentic mesh autonomously executes closed-loop interventions: triggering automated work-order generation in Computerised Maintenance Management System (CMMS), allocating spare parts inventory, reprioritising maintenance queues, and adjusting real-time haul truck dispatch trajectories to eliminate downstream crusher bottlenecks. To support robust closed-loop execution, the technical engineering team will have to enhance enterprise API connectors, transaction retry logic, and deterministic circuit breakers for safe human-agent handoffs.

Prescriptive workflows are orchestrated via AWS Step Functions and Amazon EventBridge event buses. Instead of firing static control room alerts, autonomous agents trigger automated REST API calls to external Computerised Maintenance Management Systems (CMMS) like SAP EAM to pre-order replacement spares, assign mechanics, and issue dynamic dispatch rerouting updates to Fleet Management Systems.

Stage 5: Governance Layer (Unified Operational Hub & Dynamic Oversight): A real-time operational data warehouse and analytics fabric connects frontline autonomous agent execution logs to centralised enterprise monitoring environments (e.g., Microsoft Power BI). This layer provides mine managers, reliability engineers, and executive leadership with dynamic risk auditing, continuous equipment availability metrics, ESG/carbon-intensity tracking, and full traceability over autonomous machine actions to eliminate operational blind spots and overcome the Gen AI Paradox. To guarantee enterprise compliance and observability, the technical engineering team will have to enhance automated Policy-as-Code guardrails, immutable telemetry audit logs, and unified data warehouse query performance.

Centralised data lake governance is established using Amazon Athena serverless query engines linked to an enterprise Microsoft Power BI / Amazon QuickSight hub. AWS IAM and CloudTrail enforce Policy-as-Code guardrails, providing mine management with immutable audit logs, real-time risk visibility, and carbon/ESG tracking across all autonomous agent actions.

Figure 1: The 5-Stage Closed-Loop Data Architecture Pipeline (Ingestion to Governance)



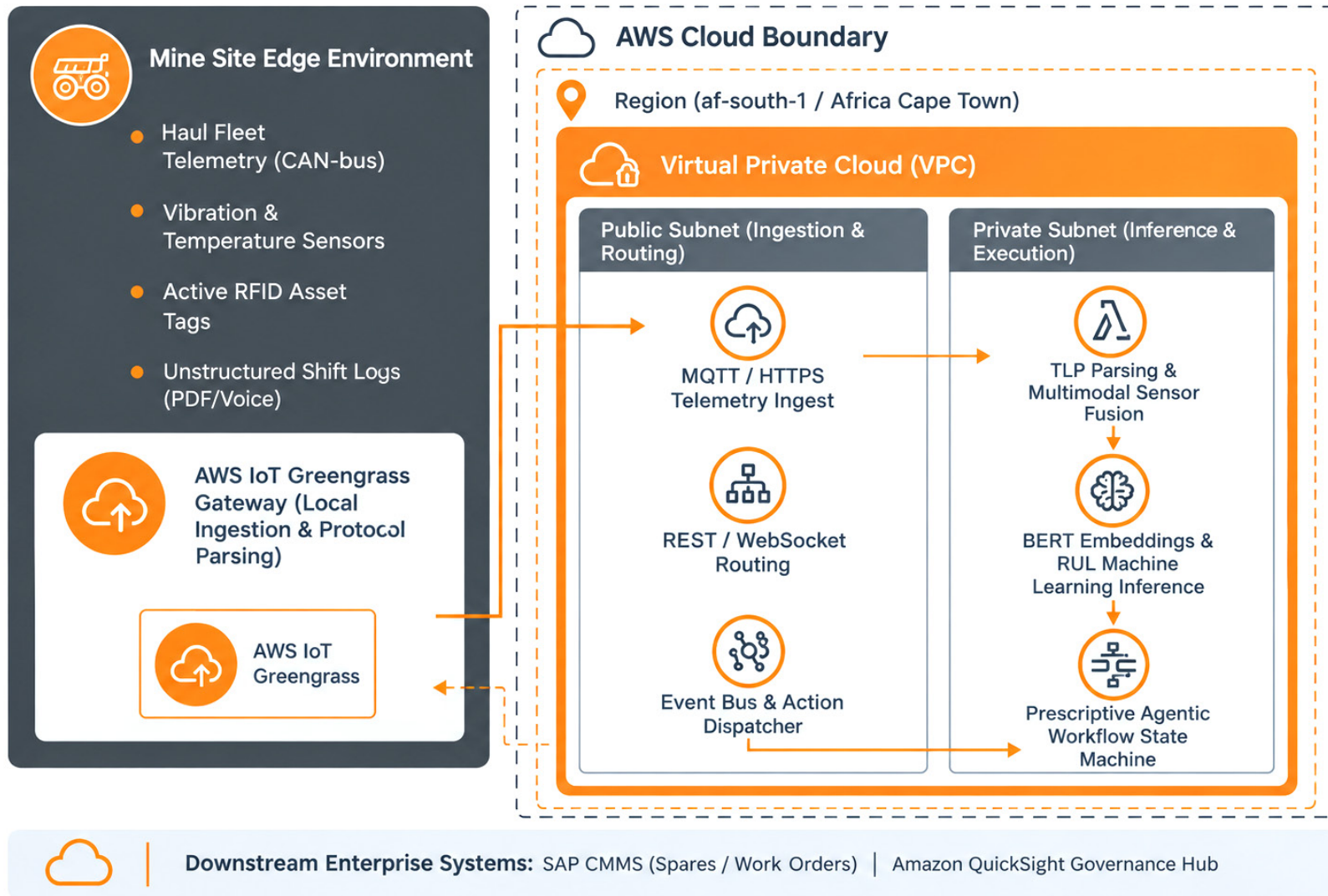
Operational Applications Across Mining & Heavy Operations

The transformation of industrial workflows follows a unified signal execution pipeline:

This closed-loop pipeline directly addresses operational friction across core mining and mineral processing domains:

Resource Sector / Application Surface Mining & Hauling Operations	Primary Field & Data Friction Decoupled weight sensors and pit maps trigger crusher queues, excessive idling, and high fuel burn. Edge Fusion & Telemetry Solution Synchronises live GNSS, road gradients, and payload data into a unified fleet state. Prescriptive AI & Decision Support AI agents dynamically reroute trucks around queues, cutting cycle delays and fuel waste.
Resource Sector / Application Heavy Equipment Fixed Plant (Crushers / Conveyors)	Primary Field & Data Friction Conflicting SCADA alarms and technician logs obscure mechanical fatigue, risking sudden plant downtime. Edge Fusion & Telemetry Solution Fuses shift notes, tri-axial vibration, and hydraulic pressures into an integrated asset health index. Prescriptive AI & Decision Support RUL models forecast failure windows while TCBR matches past cases to auto-dispatch repair orders.
Resource Sector / Application Mobile Fleet Field Service & Repairs	Primary Field & Data Friction Ambiguous OEM codes lack repair context and inventory visibility, driving repeat equipment failures. Edge Fusion & Telemetry Solution Captures degradation signals and standardises disparate OEM protocols directly on the machine. Prescriptive AI & Decision Support Mobile interfaces deliver step-by-step diagnostic fixes from historical cases and auto-reserve replacement parts.

Figure 2: AWS Reference Architecture: Pit-to-Plant Closed-Loop Agentic Operational Loop





Discussion, Business Impact & Conclusion

Discussion, Business Impact & Conclusion

Capturing raw machine telemetry or SCADA streams without prescriptive execution tools inevitably induces operational alarm fatigue; conversely, building AI dashboards without real-time sensor fusion and unstructured text abstraction produces fragile, inaccurate models. Bridging the Gen AI Paradox requires moving past horizontal copilots that offer minor 5% to 10% task-level gains. By uniting Edge IoT telemetry fusion, domain-adapted unstructured data abstraction (TLP/TCBR), and prescriptive autonomous agents, capital-intensive resource enterprises can establish closed-loop operational workflows that autonomously triage disruptions before they cascade across the value chain.

Expected Business Outcomes

- **Minimised Data Lag:** Real-time edge visibility connects distributed haul fleets, remote drill rigs, and stationary comminution circuits directly to central pit operations without telemetry blind spots.
- **Elimination of Analysis Paralysis:** Replaces alarm floods and shift-handover ambiguity with prioritised, deterministic diagnostic pathways, compressing resolution cycle times by 60% to 90% and automating up to 80% of routine triage events.
- **Unified Data Semantics:** Converts high-frequency physical telemetry, informal shift logs, and proprietary OEM formats into a standardised, queryable operational language with high entity extraction accuracy.
- **Dynamic Condition-Based Maintenance:** Shifts operations from rigid calendar schedules to predictive interventions by matching active symptoms to verified historical resolutions at 95% precision, directly extending heavy asset lifecycles.

Call to Action (CTA)

Mining and heavy industry leadership must modernize their enterprise data architectures. Escaping pilot purgatory demands deploying robust edge sensor fusion, establishing domain-adapted unstructured text abstraction layers, and constructing an auditable, real-time agentic governance hub purpose-built for physical industrial operations.