

- Clinical Operations Intelligence System
 - Executive Summary
 - The Problem
 - The Vision
 - Why This Is Different
 - Solution Architecture (High Level)
 - Agents
 - Key Capabilities Delivered
 - MVP Capabilities
 - Phase 2 Capabilities
 - Data Required
 - Implementation Phases
 - Future Expansion (not in current scope)

Clinical Operations Intelligence System

Executive Summary

The Problem

Clinical trial sponsors managing multi-site, multi-CRO studies lack unified visibility across operational signals. Enrollment shortfalls, data quality erosion, and monitoring lapses are tracked in separate vendor systems (EDC, CTMS, RBQM). Teams review these signals in isolation, missing compound risks that span domains. By the time issues surface through manual reporting cycles, they have already impacted timelines.

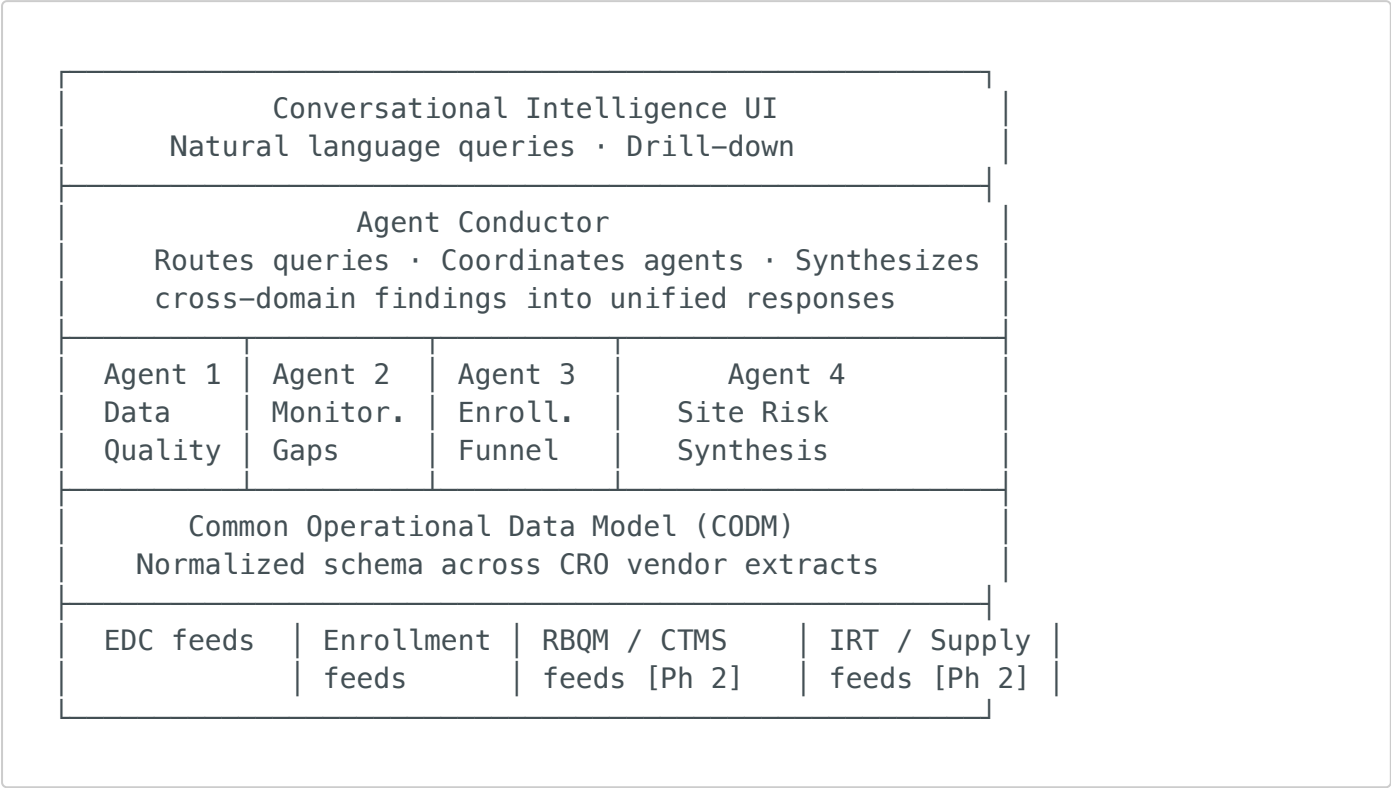
The Vision

An **AI-powered operational intelligence layer** that integrates key execution signals from CRO systems, correlates them across domains, and presents findings through a conversational interface. The system helps study teams identify which sites need attention, why, and what's compounding — replacing fragmented manual review with a unified, queryable view of trial operations.

Why This Is Different

Clinical operations teams already have dashboards — RBQM platforms surface KRIs, CTMS tracks enrollment, EDC tools flag queries. The gap is not data availability; it's **cross-domain reasoning**. No existing tool connects a site's rising query rate with its overdue monitoring visit and its declining screen-to-randomization ratio to say: "This site has a compounding problem, and here's the most likely driver." That synthesis happens in a study manager's head today, if it happens at all. This system replaces that manual pattern-matching with LLM-powered agents that reason across domains, explain their findings in natural language, and respond to follow-up questions — turning weeks-old static reports into a live investigative dialogue.

Solution Architecture (High Level)



Agents

Each agent follows a **Perceive-Reason-Plan-Act** loop — ingesting operational signals, identifying anomalies, correlating with available context, and generating findings with cited evidence.

#	Agent	What It Does
1	Data Quality	Detects query spikes, EDC entry lag, and missing data trends per site. Correlates with monitoring visit recency and site maturity (when monitoring feeds are available in Phase 2) to narrow likely drivers.
2	Monitoring Gaps	Flags sites where monitoring cadence has drifted from the risk-based plan. Prioritizes by compounding factors (overdue monitoring + deteriorating data quality).
3	Enrollment Funnel	Decomposes enrollment shortfalls by stage — screening volume, screen failure rate, randomization — to identify which constraint is binding at each site. In Phase 2, correlates with IRT supply signals to distinguish screening shortfalls from supply-driven interruptions.
4	Site Risk Synthesis	Combines signals from Agents 1-3 into a prioritized site watch list. Surfaces compound risks that no single-domain report would catch.

The **Conductor** orchestrates multi-agent queries. In MVP, it acts as a lightweight query router directing questions to Agents 1 and 3. In Phase 2, it becomes a full multi-agent orchestrator — routing cross-domain questions ("Which sites need attention and why?") to all relevant agents, collecting findings, and synthesizing a unified response.

Key Capabilities Delivered

MVP Capabilities

1. Enrollment Funnel Diagnosis — Break down enrollment shortfalls by stage (screening volume, screen failure rate, randomization) at each site. Surface *which constraint is binding* — low referral pipelines, restrictive eligibility hitting specific criteria, or operational delays — so the study team intervenes at the right point, not generically.

2. Data Quality Anomaly Investigation — Detect query spikes, EDC entry lag, and missing data trends per site. Correlate these signals with site activation recency and available operational context to narrow the likely driver and flag sites that need attention before backlog compounds.

3. Conversational Operational Q&A — Study teams ask questions in plain language ("Which sites are behind on enrollment and why?", "What's driving the query backlog at Site 042?") and get a data-grounded narrative answer drawing from available operational feeds. Follow-up questions refine the investigation without starting over.

Phase 2 Capabilities

4. Supply-Aware Enrollment Diagnosis — Enrich the enrollment funnel with IRT supply-constraint signals, allowing the agent to distinguish screening shortfalls from supply-driven interruptions (e.g., kit stockouts or randomization delays caused by depot issues). Supply-driven enrollment delays are among the most common and most misdiagnosed causes of site underperformance — teams blame sites for slow enrollment when the real cause is a depot shipping delay or kit expiry. This capability eliminates that blind spot.

5. Monitoring Gap Detection — Flag sites where monitoring cadence has drifted from the risk-based monitoring plan, prioritized by compounding factors (e.g., sites with both overdue monitoring *and* deteriorating data quality are surfaced first).

6. Cross-Domain Site Risk Prioritization — Combine signals across data quality, monitoring cadence, enrollment, and supply into a prioritized site watch list. The value is connecting dots that today sit in separate reports — a site with rising queries *and* overdue monitoring *and* slowing enrollment is a different risk profile than any one signal alone.

Data Required

Data Feed	Phase	Source Systems	Refresh	Purpose
EDC Telemetry	MVP	Medidata Rave, Oracle InForm, etc.	Daily	Entry lag, query volume, missing data, corrections
Enrollment Funnel	MVP	CTMS / IWRS	Daily	Screened/randomized counts, failure reasons

Data Feed	Phase	Source Systems	Refresh	Purpose
Monitoring Signals	Phase 2	RBQM platforms, CTMS	Weekly	Visit cadence, KRIs, overdue actions
IRT / Supply Signals	Phase 2	IRT systems (e.g., Suvoda, Signant)	Daily or event-driven	Kit inventory levels, randomization delay events, depot shipment status

All feeds are normalized into a **Common Operational Data Model (CODM)** with schema validation and feed-health monitoring at ingestion. Stale or missing feeds trigger explicit data currency warnings on all downstream outputs.

The two MVP data feeds (EDC, Enrollment) are standard CRO deliverables with established flat-file export mechanisms, minimizing integration risk for initial deployment. Phase 2 adds Monitoring and IRT / Supply feeds. IRT integration may require API-level access or event-driven pipelines rather than flat-file exports, which is a different integration pattern — the phased approach gives the team time to negotiate data access and build the appropriate connectors.

A note on CODM normalization: EDC schemas vary across vendor platforms (Rave, InForm, Veeva). Mapping these into a common model is non-trivial and is expected to be the primary technical effort in the data layer. The CODM design should be protocol-aware and version-controlled, with explicit mapping documentation per source system.

Implementation Phases

Phase	Scope
1 — MVP	Agents 1 & 3 (Data Quality + Enrollment Funnel) · EDC & enrollment feeds · Lightweight query router · Conversational Q&A interface
2 — Full Scope	Add Agent 2 (Monitoring Gaps) + Agent 4 (Site Risk Synthesis) · RBQM/CTMS monitoring feeds · IRT / Supply feed integration · Supply-context enrichment for Enrollment Funnel agent · Full Conductor multi-agent orchestration · Cross-domain site risk prioritization

Future Expansion (not in current scope)

Additional domains (milestone tracking, vendor SLA monitoring, portfolio-level intelligence) can be added as new agents once the core platform is proven and the required data integrations (vendor portals, cross-study data) are established. These are deliberately excluded from current scope due to higher data access barriers and lower feasibility confidence.