



Daram Partheev,
IIM Indore,
B.A. + MBA



Harisharnam,
IIM Indore,
B.A. + MBA



Sai Sathwik Pabba,
NIT Bhopal,
B.tech (ECE)



Kritika,
Jamia Humdard,
B.Pharm

An AI-Assisted Clinical Trial Operations Control System

Making risk visible, decisions explainable, and actions auditable

Problem Statement 1: Integrated insight-driven data-flow model

Critical Risk Detection

Data Governance

Action Readiness

Team: Pharma Flux

Mentor: Prof. Aditya Maheshwari
PhD (IIT Bombay), Faculty IIM Indore

Executive Overview

⚠ The Problem

- Clinical trials generate siloed operational and clinical data.
- Monitoring remains manual, reactive, and often delayed by weeks.
- Existing tools lack explainability and unified governance.

📦 Our Solution: NEST 2.0

- An end-to-end control system converting raw data into actionable risk signals.
- **Workflow:** Data Ingest → Signal Detection → Governed Actions → AI Insights.
- Features human-in-the-loop validation for every AI recommendation.

↗ Key Outcomes

- **Faster Intervention:** Detect risks days earlier, not weeks later.
- **Reduced CRA Burden:** AI drafts actions, humans approve.
- **Data Trust:** Full audit trails and explainable logic.

TARGET IMPACT

100% Audit Readiness

Why Current Clinical Monitoring Fails

⌚ Reactive & Fragmented

- Review cycles are retrospective, causing weeks of delay in risk detection.
- Dashboards are static, offering lagging indicators instead of real-time signals.

劳累 Operational Overload

- CRAs are overwhelmed by manual data triangulation across disparate systems.
- **No Closed Loop:** Insights are generated but lack a tracked, governed action path.

locker Governance Gaps

- Current tools lack transparency ("Black Box" algorithms).
- No audit-safe mechanism to validate AI-driven decisions for inspections.

“



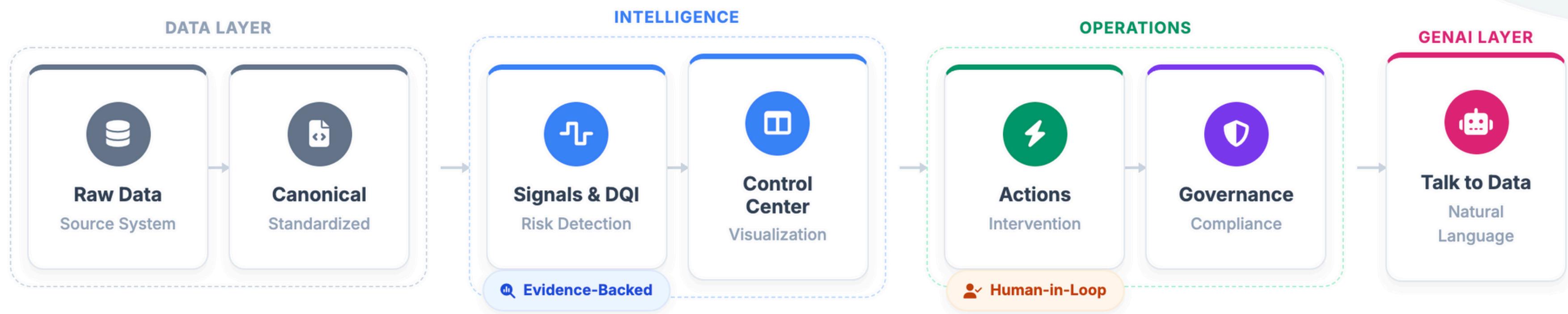
The Missing Link

"There is no single system today that effectively links Data → Insight → Action → Governance."

CRITICAL OPERATIONAL GAP

Solution at a Glance

End-to-End Clinical Data Operations Pipeline



Control Center — Real-time Oversight

- **Unified Operational View**

Provides a single pane of glass for cross-study operational health, replacing fragmented spreadsheets.

Clinical Ops

- Quick Create
- Control Center
- Talk to Data
- Sites
- Actions
- Governance
- Documents
- Data Library
- More

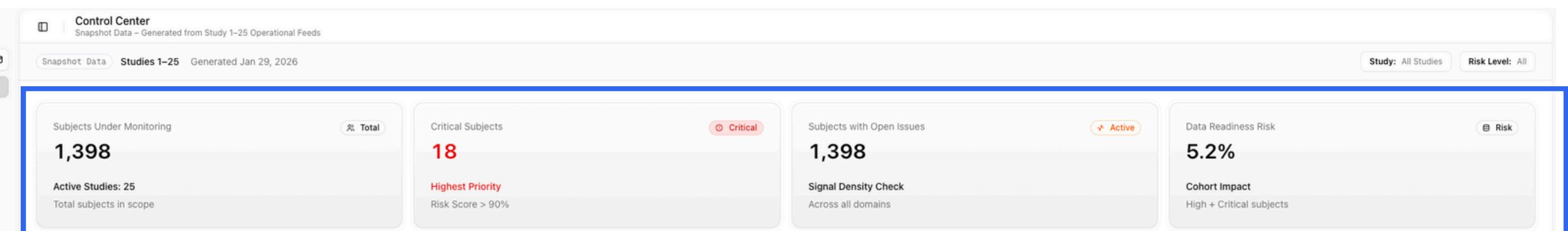
Settings

Search

CN Dr. Analyst analyst@novartis.com

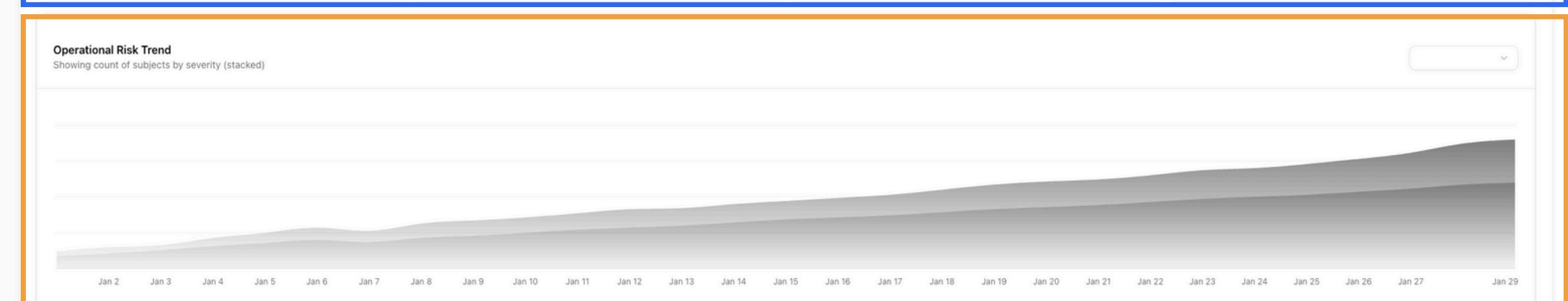
- **Global KPIs & Scale**

Instantly visualize monitoring scope (1,398 subjects) and critical data readiness risks.



- **Risk Trend Analysis**

Tracks severity over time to identify emerging operational bottlenecks before they escalate.



- **Prioritized Actions**

Directs CRAs to high-risk subjects requiring immediate intervention.

| Subject ID | Study | Primary Risk Domain | Risk Level | DQI Score | Signal Count | Assigned CRA |
|------------|----------|---------------------|------------|-----------|--------------|--------------|
| SUB-6133 | STUDY-23 | EDC Completeness | Critical | 99.30 | 15 | Unassigned |
| SUB-4886 | STUDY-19 | EDC Completeness | Critical | 98.20 | 13 | Unassigned |
| SUB-47 | STUDY-1 | EDC Completeness | Critical | 97.90 | 13 | Unassigned |
| SUB-83 | STUDY-1 | EDC Completeness | Critical | 97.50 | 18 | Unassigned |
| SUB-13835 | STUDY-24 | EDC Completeness | Critical | 97.40 | 11 | Unassigned |

Global KPI's

Risk Trend

Prioritised Subjects

KPI Summary Cards

- **Subjects Under Monitoring**

Total scope of 1,398 subjects across 25 active studies, establishing the scale of operational oversight.

- **Open Issues**

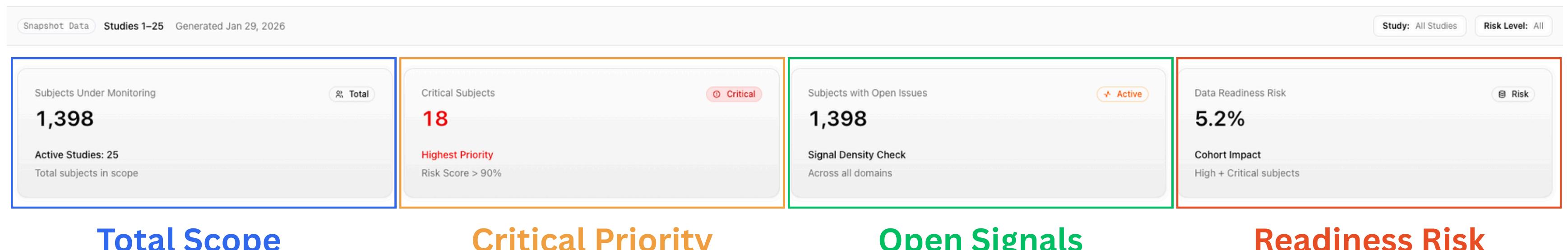
Tracks subjects with active signal density checks across all domains, highlighting workflow bottlenecks.

- **Critical Subjects**

Identifies 18 high-priority subjects with Risk Score > 90%, enabling immediate CRA triage.

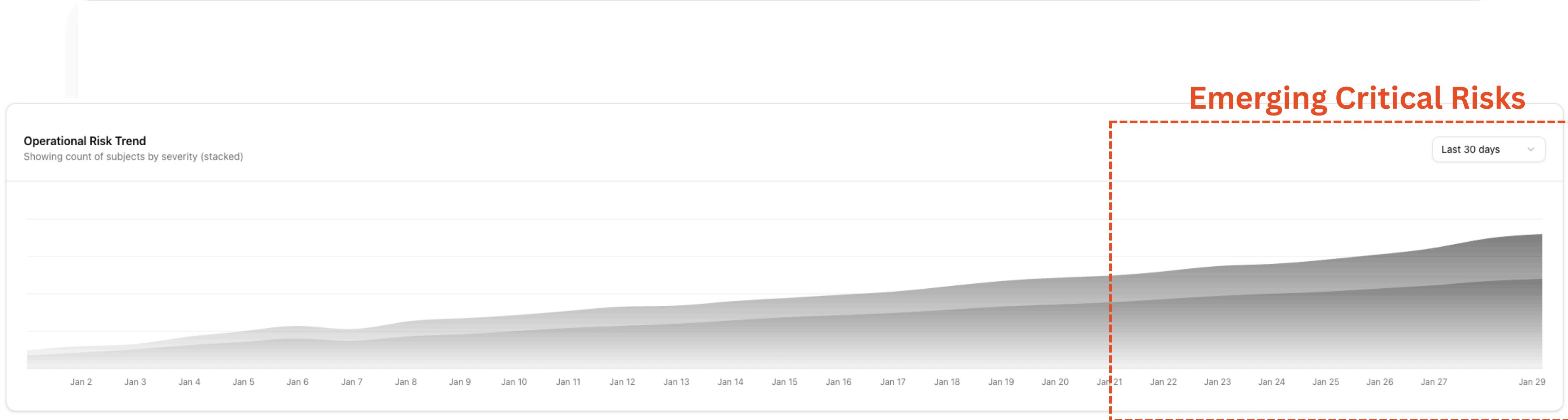
- **Data Readiness Risk**

Quantifies cohort impact (5.2%) of high + critical subjects, providing a direct metric for data trust.



Operational Risk Trend

Longitudinal view of emerging risks across study cohorts (30-Day View)



Trend-Based Monitoring

Shifts focus from static "snapshot" metrics to dynamic risk velocity. Detects when issues are accumulating faster than resolution rates.



Early Warning System

Upward trajectory in the stacked chart serves as an early warning for site saturation or systemic protocol deviations before they become critical.



Severity Stratification

Visualizes the proportion of critical vs. moderate risks (dark vs. light grey areas), allowing teams to prioritize "red" zones immediately.



Actionable Risk Table

Subject-level prioritization, data quality scoring, and resource assignment

Outline Domain Analysis 2 CRA Assignments 1 Risk Overview 4

| | | Study | Primary Risk Domain | Prioritisation | Data Quality | | |
|--------------------------|------------|----------|---------------------|----------------|--------------|--------------|--------------|
| | | | | Risk Level | DQI Score | Signal Count | Assigned CRA |
| <input type="checkbox"/> | Subject ID | STUDY-20 | EDC Completeness | High | 58.30 | 4 | Unassigned |
| <input type="checkbox"/> | SUB-2233 | STUDY-10 | EDC Completeness | Low | 15.50 | 1 | Unassigned |
| <input type="checkbox"/> | SUB-11839 | STUDY-21 | EDC Completeness | Medium | 21.40 | 1 | Unassigned |
| <input type="checkbox"/> | SUB-4355 | STUDY-16 | EDC Completeness | Medium | 21.40 | 1 | Unassigned |
| <input type="checkbox"/> | SUB-37812 | STUDY-22 | EDC Completeness | Medium | 21.40 | 1 | Unassigned |
| <input type="checkbox"/> | SUB-18683 | STUDY-21 | EDC Completeness | Medium | 21.40 | 1 | Unassigned |
| <input type="checkbox"/> | SUB-3670 | STUDY-19 | EDC Completeness | Medium | 21.40 | 1 | Unassigned |
| <input type="checkbox"/> | SUB-31881 | STUDY-22 | EDC Completeness | Medium | 21.40 | 1 | Unassigned |
| <input type="checkbox"/> | SUB-36810 | STUDY-22 | EDC Completeness | Medium | 21.40 | 1 | Unassigned |



Risk-Based Prioritization

Subjects are automatically sorted by risk severity (Critical/High), ensuring teams focus on the most urgent operational issues first.



DQI Integration

Real-time Data Quality Index (DQI) scores (e.g., 99.30%) provide immediate context on the reliability of the subject's dataset.



Clear Ownership

Direct CRA assignment and "Unassigned" flagging prevents critical subjects from slipping through the cracks during monitoring.

Click to see

Actionable Risk Table

Subject-level prioritization, data quality scoring, and resource assignment

Subject wise Active Signals

The screenshot displays the Actionable Risk Table interface. At the top, there are four cards: 'Active Studies: 25' (Total subjects in scope), 'Highest Priority' (Risk Score > 90%), 'Signal Density Check' (Across all domains), and 'Cohort Impact' (High + Critical Subjects). Below these is a 'Operational Risk Trend' chart showing the count of subjects by severity from Jan 2 to Jan 24. The main area is a table with columns: Subject ID, Study, Primary Risk Domain, Risk Level, DQI Score, Signal Count, and Assigned CRA. The table lists subjects like SUB-6133, STUDY-23, EDC Completeness, Critical, 99.30, 15, Unassigned. A red dashed box highlights the 'SUB-6133' row. To the right of this box is a detailed view for 'SUB-6133' with a DQI score of 99.30 (Critical Risk). This view includes a 'Risk Contribution by Domain' chart (EDC Completeness: 67% Contrib., Visit Compliance: 33% Contrib.) and a list of 'Active Signals (Grouped)' under 'EDC COMPLETENESS (10 ISSUES)'. Each signal is a card with a 'View Evidence' button. A large black arrow points from the 'Assigned CRA' column towards the 'View Evidence' button for the first signal. At the bottom right of this view are 'Save Review' and 'Close' buttons.

| Subject ID | Study | Primary Risk Domain | Risk Level | DQI Score | Signal Count | Assigned CRA |
|------------|----------|---------------------|------------|-----------|--------------|--------------|
| SUB-6133 | STUDY-23 | EDC Completeness | Critical | 99.30 | 15 | Unassigned |
| SUB-4886 | STUDY-19 | EDC Completeness | Critical | 98.20 | 13 | Unassigned |
| SUB-47 | STUDY-1 | EDC Completeness | Critical | 97.90 | 13 | Unassigned |
| SUB-83 | STUDY-1 | EDC Completeness | Critical | 97.50 | 18 | Unassigned |
| SUB-13835 | STUDY-24 | EDC Completeness | Critical | 97.40 | 11 | Unassigned |
| SUB-9118 | STUDY-24 | EDC Completeness | Critical | 97.20 | 13 | Unassigned |
| SUB-3567 | STUDY-24 | EDC Completeness | Critical | 96.60 | 13 | Unassigned |
| SUB-9302 | STUDY-24 | EDC Completeness | Critical | 96.30 | 10 | Unassigned |
| SUB-12 | STUDY-1 | EDC Completeness | Critical | 95.80 | 14 | Unassigned |

Risk-Based Prioritization

Subjects are automatically sorted by risk severity (Critical/High), ensuring teams focus on the most urgent operational issues first.

DQI Integration

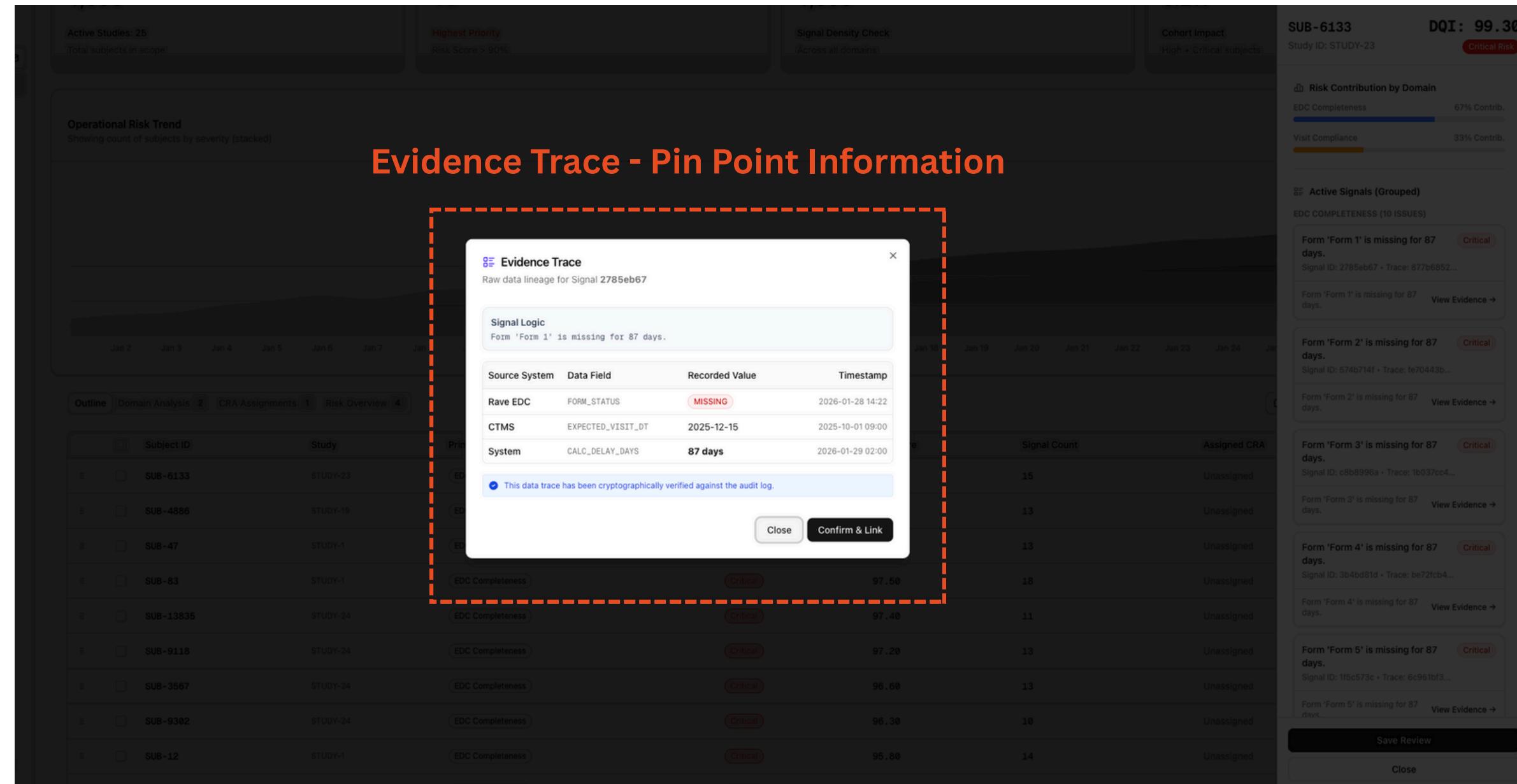
Real-time Data Quality Index (DQI) scores (e.g., 99.30%) provide immediate context on the reliability of the subject's dataset.

Clear Ownership

Direct CRA assignment and "Unassigned" flagging prevents critical subjects from slipping through the cracks during monitoring.

Actionable Risk Table

Subject-level prioritization, data quality scoring, and resource assignment



Evidence Trace - Pin Point Information

A modal window titled "Evidence Trace" displays raw data lineage for Signal 2785eb67. It shows Signal Logic: "Form 'Form 1' is missing for 87 days." Below this is a table of source systems, data fields, recorded values, and timestamps. A note at the bottom states: "This data trace has been cryptographically verified against the audit log." Buttons for "Close" and "Confirm & Link" are at the bottom right.

| Source System | Data Field | Recorded Value | Timestamp |
|---------------|-------------------|----------------|------------------|
| Rave EDC | FORM_STATUS | MISSING | 2026-01-28 14:22 |
| CTMS | EXPECTED_VISIT_DT | 2025-12-15 | 2025-10-01 09:00 |
| System | CALC_DELAY_DAYS | 87 days | 2026-01-29 02:00 |



Risk-Based Prioritization

Subjects are automatically sorted by risk severity (Critical/High), ensuring teams focus on the most urgent operational issues first.



DQI Integration

Real-time Data Quality Index (DQI) scores (e.g., 99.30%) provide immediate context on the reliability of the subject's dataset.

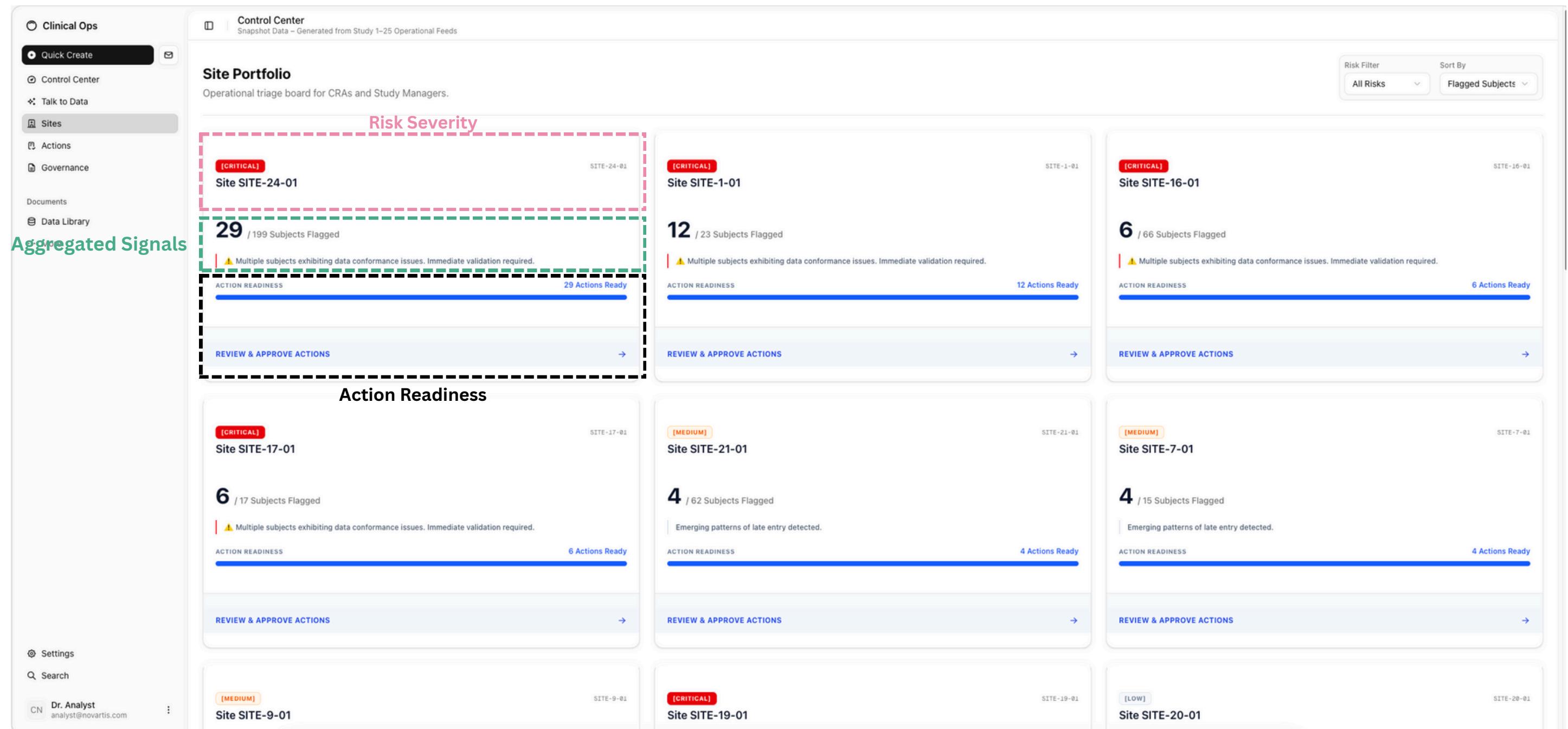


Clear Ownership

Direct CRA assignment and "Unassigned" flagging prevents critical subjects from slipping through the cracks during monitoring.

Sites — Portfolio View

Operational triage board for CRAs and Study Managers



Site-Level Aggregation

Risks are not isolated events. The system aggregates individual subject signals to calculate a composite risk score for the entire site.



Risk Density Check

Visualizes the proportion of flagged subjects (e.g., 29/199) to distinguish between isolated incidents and systemic site failure.



Immediate Validation

"Action Readiness" bars indicate where AI has already drafted interventions, allowing managers to prioritize sites with pending decisions.

Site Card Anatomy

- **Instant Risk Triage**

The "Critical" badge instantly flags high-risk sites, allowing CRAs to prioritize interventions efficiently within a busy portfolio.

- **Issue Density Signal**

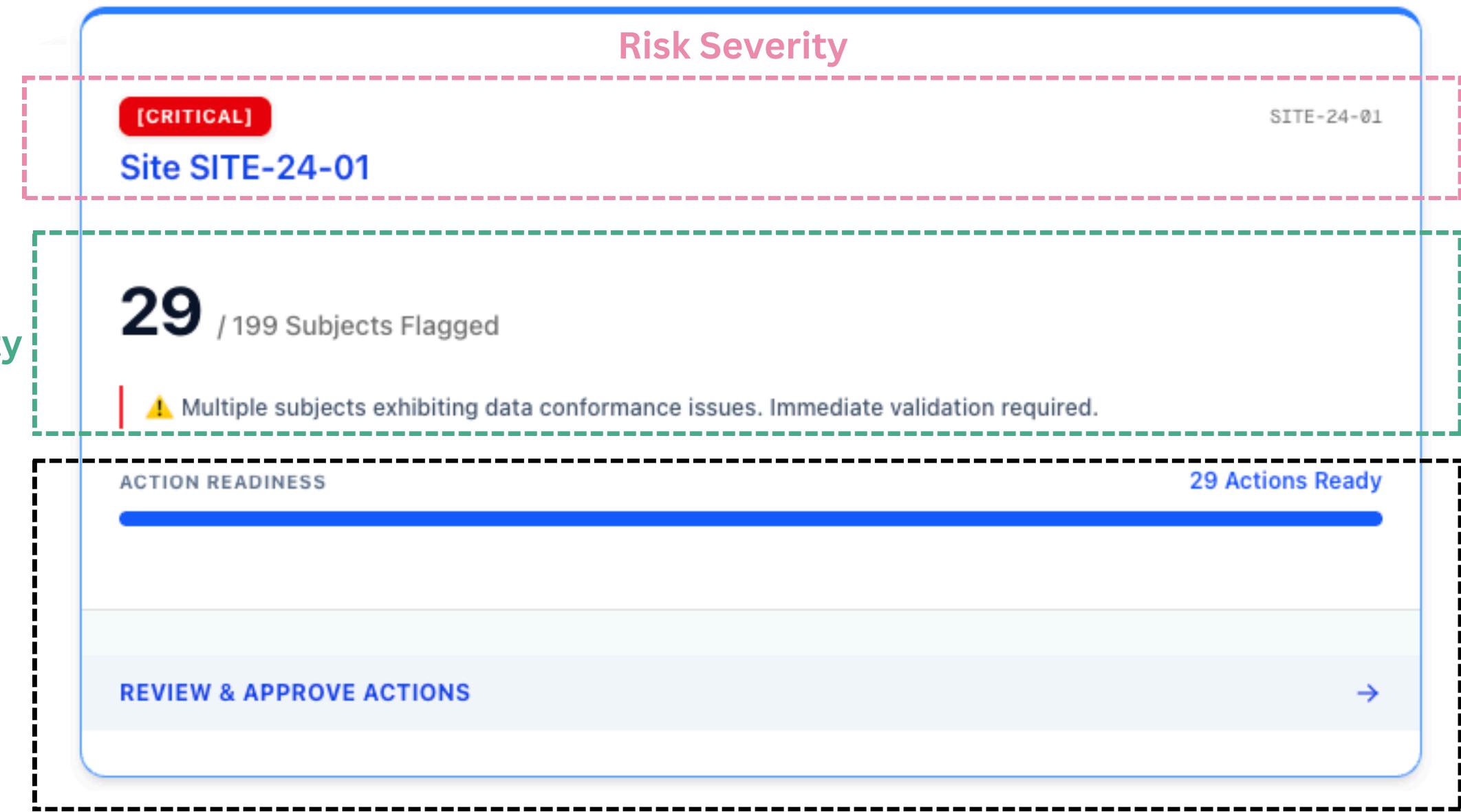
Metrics like "29/199 Subjects Flagged" provide immediate context on the scale of data conformance issues versus total enrollment. **Flagged Density**

- **Action-Oriented Workflow**

The readiness bar aggregates 29 drafted actions, enabling rapid bulk-processing rather than manual, row-by-row investigation.

- **Immediate Validation**

Visual warnings highlight specific conformance patterns, prompting immediate human-in-the-loop validation.



Action Readiness

Subject Drill-Down (Site 24-01) (Example)

Root cause analysis with explainable AI narrative and evidence tracing

DEEP DIVE ANALYSIS

Site SITE-24-01 (SITE-24-01) Critical Risk
199 Subjects Enrolled 29 Flagged 29 Pending Actions

SELECT SUBJECT FOR ANALYSIS: SUB-13835 (Flagged)

Operational Insight for SUB-13835
AI analysis indicates data anomalies consistent with other subjects at this site. Review the evidence below before approving actions.

AI Operational Analysis
Automated analysis of signal patterns for Subject SUB-13835

AI Narrative & Pattern

Recommended Actions

- Email to Site Coordinator** (High)
Draft an email to the site coordinator to address the missing forms for subject SUB-13835.
Subject: Urgent: Missing Forms for Subject SUB-13835
Dear [Site Coordinator's Name],
I hope this message finds you well. We have identified a critical issue regarding the electronic data capture (EDC) completeness for subject SUB-13835. There are several missing forms (Form 12, 19, 21, 22, 23) which have been missing for 58 days, indicating a significant gap in data completeness that requires immediate attention.
Expected Impact: Reduce open visit queries by ~20% (Est.)
Why this action?
Edit Reject Approve
- Review Site Training Records** (Medium)
Create a task to review the training records of the site staff to identify any gaps that might have contributed to the missing forms.
Review the training records of the site staff responsible for data entry for subject SUB-13835. Check for any gaps or recent changes in staff that might have led to the missing forms. Ensure that all staff are adequately trained on the EDC system and the importance of timely data entry.
Expected Impact: Clarify protocol deviation within 24h
Why this action?
Edit Reject Approve

Pending Action

Pattern Confidence

| | | |
|-----------------------|------------------------|---------------|
| ACTION READINESS HIGH | PATTERN CONFIDENCE 89% | EVIDENCE 100% |
|-----------------------|------------------------|---------------|

Generative Explanation

The AI doesn't just flag "Site Training Issues"; it synthesizes a natural language narrative explaining **why**, citing specific missed visits and protocol deviations (Conf: 88%).

Evidence & Confidence

Every insight is scored. **89% Pattern Confidence** and **100% Evidence Coverage** ensure CRAs know exactly how reliable the signal is before acting.

Instant Remediation

The system drafts a relevant "Follow-up on Outstanding Visit" email immediately, closing the gap between insight and intervention.

AI Operational Analysis (Site 8-01) (Example)

Automated signal pattern detection and root cause hypothesis generation

CAUSAL ANALYSIS

[← Back to Site Portfolio](#)

Site SITE-8-01 (SITE-8-01) Low Risk
37 Subjects Enrolled 0 Flagged 0 Pending Actions

SELECT SUBJECT FOR ANALYSIS: SUB-911

Action Readiness

| | | |
|---------------------------------|---------------------------|------------------|
| ACTION READINESS HIGH | PATTERN CONFIDENCE 89% | EVIDENCE 100% |
|---------------------------------|---------------------------|------------------|

Operational Insight for SUB-911
AI analysis indicates data anomalies consistent with other subjects at this site. Review the evidence below before approving actions.

AI Operational Analysis
Automated analysis of signal patterns for Subject SUB-911

Narrative
Site training issue
The critical delay in the 'Week120' visit suggests that site staff may not be adequately trained on the importance of timely visit scheduling and follow-up. Training issues can lead to misunderstandings about protocol timelines and priorities, resulting in missed or delayed visits.
Conf: 80% IDs: 0

Pattern
Staffing shortages
A delay of 29 days in a critical visit could be due to insufficient staffing at the site, which can lead to scheduling backlogs and prioritization issues. Staffing shortages are a common operational issue that can impact visit compliance.
Conf: 70% IDs: 0

Hypotheses
Connectivity issues
If the site relies on electronic systems for scheduling and reminders, connectivity issues could prevent timely notifications and updates, leading to missed visits. This is particularly relevant if the site is in a location with unreliable internet access.
Conf: 60% IDs: 0

AI-Generated Hypotheses

Recommended Actions

Select All Drafts [Review Required](#)

EMAIL Follow-up on Outstanding Visit High
Draft an email to the site coordinator to address the outstanding visit for subject SUB-911.
Subject: Urgent: Outstanding Visit for Subject SUB-911
Dear [Site Coordinator's Name].
I hope this message finds you well. I am writing to bring to your attention that the visit scheduled for 'Week120' for subject SUB-911 has
Expected Impact: Reduce open visit queries by ~20% (Est.)
Why this action? [Edit](#) [Reject](#) [Approve](#)

TASK Review Site Capacity and Training Medium
Assign a task to review the site's capacity and training records to identify potential overload or training gaps.
Task: Review the site's capacity and training records for potential overload or training gaps that may be contributing to the delay in visit compliance for subject SUB-911. This should include an assessment of current workload, staffing levels, and recent training sessions attended by site staff.
Expected Impact: Clarify protocol deviation within 24h
Why this action? [Edit](#) [Reject](#) [Approve](#)

Explainable "Why"

The system moves beyond simple flagging. By toggling between **Narrative**, **Pattern**, and **Hypotheses**, CRAs gain a complete context of the operational risk profile.

Root Cause Identification

AI automatically suggests probable causes such as "**Site training issue**" (88% conf) or "**Staffing shortages**", directing investigators to the source of the problem.

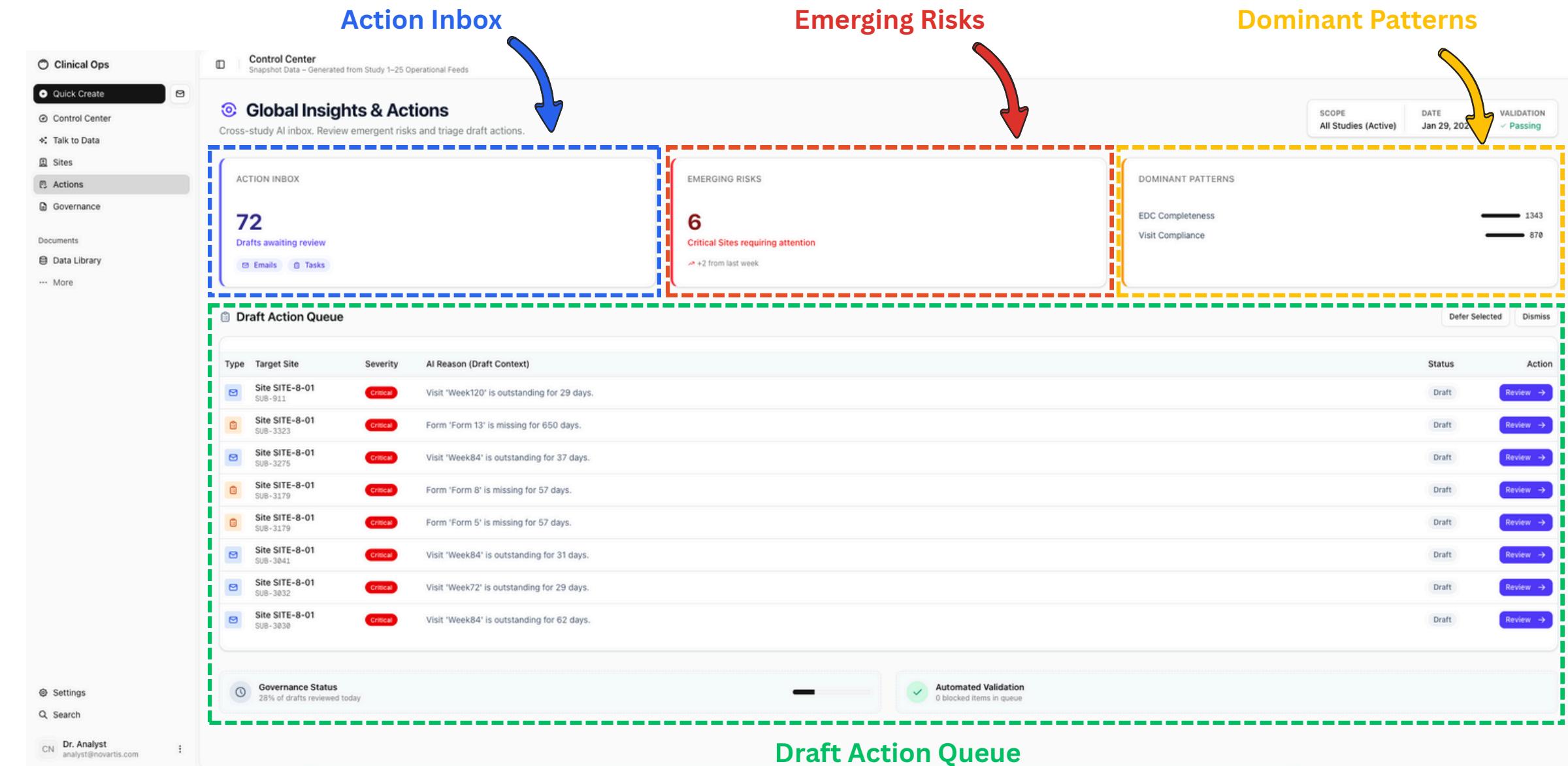
Actionable Confidence

High pattern confidence (89%) and evidence coverage translate directly to **High Action Readiness**, enabling safer, faster, and more decisive interventions.

Actions — Close the Loop

- Action Inbox & Triage**

Centralized hub for all pending AI-drafted actions, eliminating email chaos.



- Emerging Risks**

Highlighting critical sites requiring immediate attention before they fail inspection.

- Dominant Patterns**

AI identifies systemic issues (e.g., EDC Completeness) across multiple sites.

- From Signal to Action**

Converts raw risk signals into governed, trackable workflows at scale.

Action Review and Approval

Governed decision-making with predictive impact analysis

✓ Human-in-the-Loop

Recommended Actions

Select All Drafts

AI-generated Draft Context

EMAIL Email to Site Coordinator
Draft an email to the site coordinator to address the missing forms for subject SUB-13835.

Subject: Urgent: Missing Forms for Subject SUB-13835

Dear [Site Coordinator's Name],

I hope this message finds you well. We have identified a critical issue regarding the electronic data capture (EDC) completeness for subject SUB-13835. There are several missing forms across various visits, which may impact the quality of the data. I am reaching out to request your immediate attention to this matter.

Outcome Prediction

Expected Impact: Reduce open visit queries by ~20% (Est.)

Why this action?

Detailed Evidence Based-Action

Human Control

 Edit Reject Approve

TASK Review Site Training Records
Create a task to review the training records of the site staff to identify any gaps that might have contributed to the missing forms.

Review the training records of the site staff responsible for data entry for subject SUB-13835. Check for any gaps or recent changes in staff that might have led to the missing forms. Ensure that all staff are adequately trained on the EDC system and the importance of timely data entry.

Expected Impact: Clarify protocol deviation within 24h

Why this action?

Edit Reject Approve

Review Required

Measurable Impact



Before approving, CRAs see the predicted outcome (e.g., "Reduce open visit queries by ~20%"). This ensures resources are allocated to high-ROI actions.

Decision Rights



The "Approve / Reject" workflow acts as a hard governance gate. AI drafts the email or task, but the CRA retains final authority to edit the tone or defer the action.

Evidence-Backed

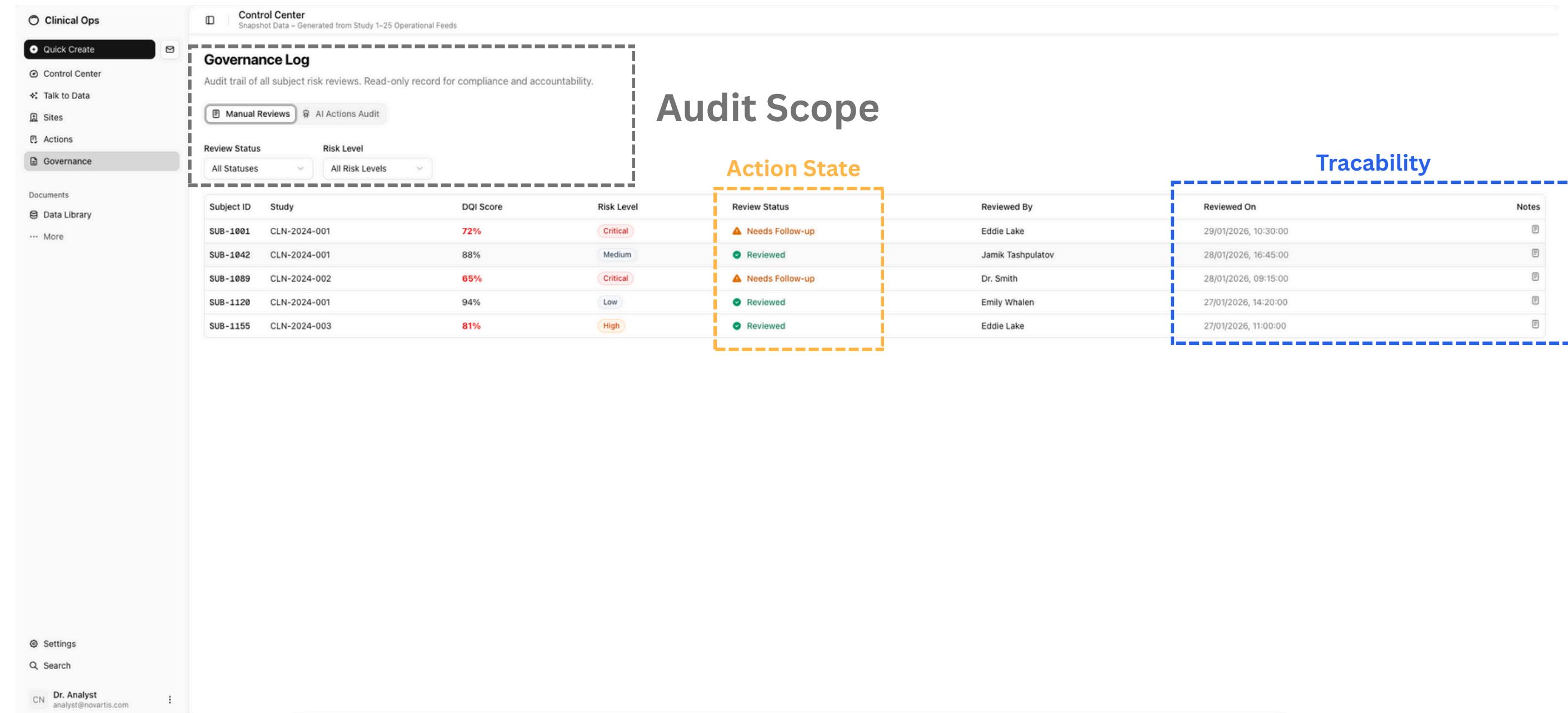


The "Why this action?" toggle provides instant access to the raw data lineage, giving users confidence that the recommendation is not a hallucination.

Governance — Inspection Readiness

End-to-end accountability and auditability across all study actions

 21 CFR Part 11 Compliant



The screenshot shows the Novartis Clinical Ops platform interface. On the left, a sidebar menu includes Clinical Ops, Quick Create, Control Center, Talk to Data, Sites, Actions, and **Governance**, which is currently selected. Below this are Documents, Data Library, and More. At the bottom are Settings, Search, and a user profile for Dr. Analyst.

The main area displays the **Control Center** with a **Snapshot Data – Generated from Study 1–25 Operational Feeds**. A **Governance Log** section is highlighted with a dashed orange border. It contains a table of subject IDs, studies, DQI scores, and risk levels. The table shows:

| Subject ID | Study | DQI Score | Risk Level |
|------------|--------------|-----------|------------|
| SUB-1001 | CLN-2024-001 | 72% | Critical |
| SUB-1042 | CLN-2024-001 | 88% | Medium |
| SUB-1089 | CLN-2024-002 | 65% | Critical |
| SUB-1120 | CLN-2024-001 | 94% | Low |
| SUB-1155 | CLN-2024-003 | 81% | High |

A **Action State** section is also highlighted with a dashed orange border. It lists review status categories: Needs Follow-up (orange triangle) and Reviewed (green circle). The table shows:

| Review Status | Reviewed By |
|-------------------|-------------------|
| ▲ Needs Follow-up | Eddie Lake |
| ● Reviewed | Jamik Tashpulatov |
| ▲ Needs Follow-up | Dr. Smith |
| ● Reviewed | Emily Whalen |
| ● Reviewed | Eddie Lake |

A **Tracability** section is highlighted with a dashed blue border. It lists reviewed dates and users:

| Reviewed On | Notes |
|----------------------|-------|
| 29/01/2026, 10:30:00 | |
| 28/01/2026, 16:45:00 | |
| 28/01/2026, 09:15:00 | |
| 27/01/2026, 14:20:00 | |
| 27/01/2026, 11:00:00 | |

Scope Control

Granular filters allow auditors to slice logs by risk level, study, or review status, instantly retrieving evidence for inspection inquiries.

Closed-Loop Tracking

Every risk signal is tracked from "Initiated" to "Resolved," preventing open issues from lingering unnoticed in the system.

Immutable Attribution

The system automatically records the **Who** (User), **When** (Timestamp), and **What** (Decision) for every action, ensuring full data integrity.

Audit Trail Details

Granular, immutable record of every system interaction for full traceability

Read-Only Record

Governance Log

Audit trail of all subject risk reviews. Read-only record for compliance and accountability.

Manual Reviews AI Actions Audit

| Timestamp | User | Type | Detailed Action | Target |
|----------------------|----------------------|-------|------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| 29/01/2026, 16:50:30 | Eddie Lake (AI Lead) | query | {"type": "query", "title": "Query Site for Visit Scheduling Challenges", "description": "Initiate a query to the site to understand specific ch..."} | SUB-911 |
| 29/01/2026, 16:50:29 | Eddie Lake (AI Lead) | task | {"type": "task", "title": "Review Site Capacity and Training", "description": "Assign a task to review the site's capacity and training adequa..."} | SUB-911 |
| 29/01/2026, 16:50:25 | Eddie Lake (AI Lead) | email | {"type": "email", "title": "Follow-up on Outstanding Visit", "description": "Draft an email to the site coordinator to address the outstanding..."} | SUB-911 |

Identity & Time

Detailed Action

State



Temporal Precision

Timestamps are captured to the exact second for every event, providing an indisputable chronological sequence for inspection reconstruction.



Action Attribution

Every log entry is strictly linked to a verified user role (e.g., AI Lead) and specific action type, ensuring clear accountability for all changes.



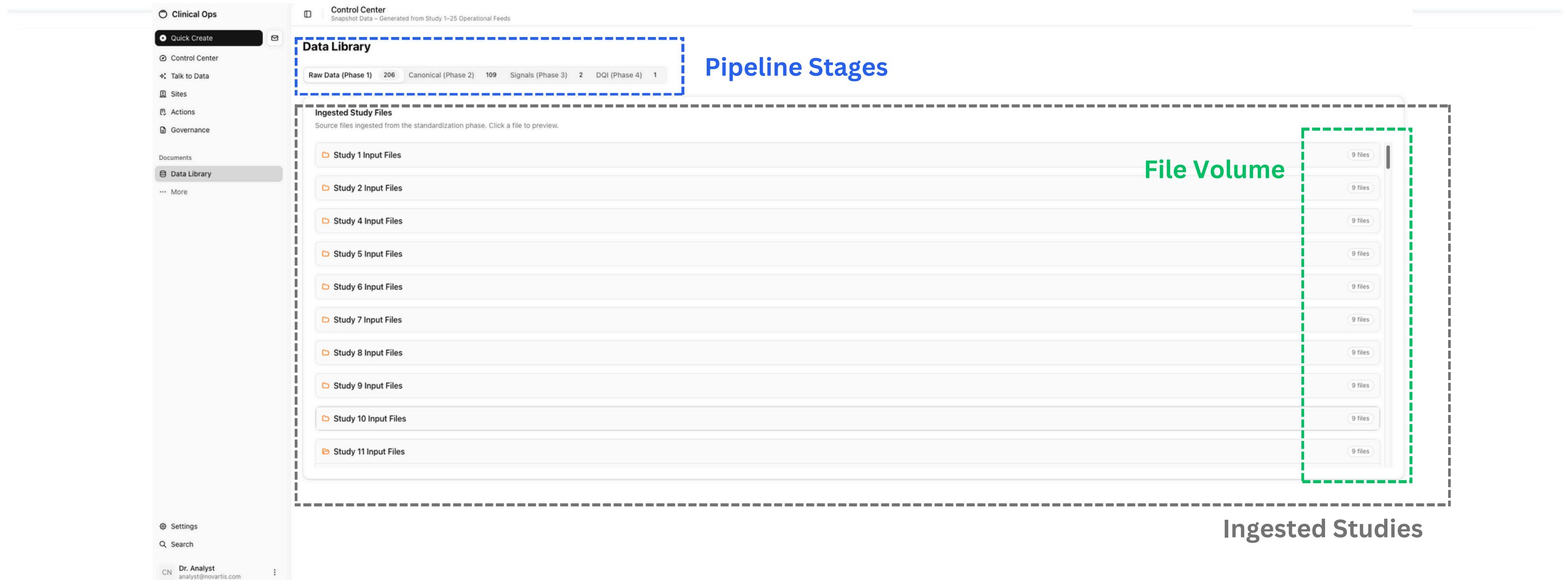
Lifecycle Transparency

The status column provides immediate visibility into the current state of every governance action, from 'Initiated' to final resolution.

Data Library — Trust Before Intelligence

Curated, versioned inputs for reproducible clinical analytics

 Source of Truth



The screenshot shows the Novartis Data Library interface. On the left is a sidebar with 'Clinical Ops' and 'Data Library' selected. The main area has a 'Control Center' header and a 'Data Library' section. The 'Data Library' section is highlighted with a blue dashed box and contains a 'Pipeline Stages' heading. Below it is a 'Ingested Study Files' section with a list of study input files from Study 1 to Study 11. To the right, a large dashed box covers the main content area. Inside this box, the word 'File Volume' is in green at the top, and below it is a vertical column of 11 boxes, each labeled '9 files'. At the bottom right of the main area, the text 'Ingested Studies' is visible.

Pipeline Stages

Ingested Study Files

- Study 1 Input Files
- Study 2 Input Files
- Study 4 Input Files
- Study 5 Input Files
- Study 6 Input Files
- Study 7 Input Files
- Study 8 Input Files
- Study 9 Input Files
- Study 10 Input Files
- Study 11 Input Files

File Volume

Ingested Studies



Controlled Pipeline

Data moves through strict phases: **Raw** → **Canonical** → **Signals** → **DQI**. This isolation ensures that analytics are always run on clean, governed datasets.



Structured Ingestion

The library organizes disparate inputs (EDC, Labs, CTMS) into study-specific containers, creating a unified directory for all operational data.



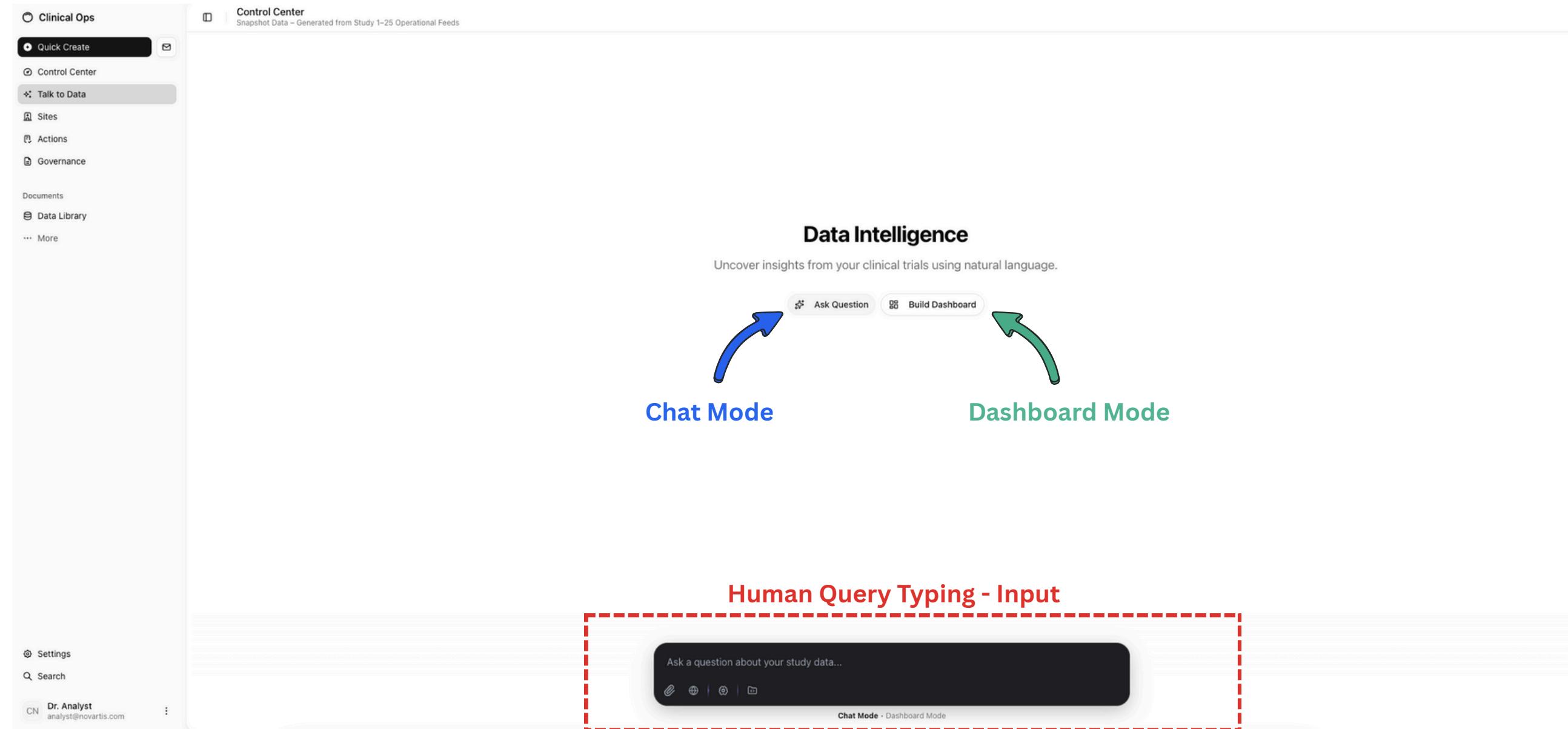
Reproducibility

Every file version is tracked with timestamps. If an insight is questioned during an inspection, the exact source file state can be retrieved.

Talk to Data — Natural Language Access

Empowering clinical teams to ask questions, not write SQL

Generative AI Powered



The screenshot shows the Novartis Clinical Ops Control Center. On the left, a sidebar menu includes 'Clinical Ops' (Quick Create, Control Center, Talk to Data), 'Sites', 'Actions', 'Governance', 'Documents', 'Data Library', and 'More'. The main area is titled 'Control Center' with the sub-tile 'Snapshot Data - Generated from Study 1-25 Operational Feeds'. Below this is a section titled 'Data Intelligence' with the sub-tile 'Uncover insights from your clinical trials using natural language.' It features two buttons: 'Ask Question' (blue) and 'Build Dashboard' (green). A blue curved arrow labeled 'Chat Mode' points to the 'Ask Question' button, and a green curved arrow labeled 'Dashboard Mode' points to the 'Build Dashboard' button. At the bottom of the main area is a red dashed box containing a black input field with the placeholder 'Ask a question about your study data...' and a small toolbar below it.

Conversational Analytics

Study Managers can interrogate data in plain English (e.g., "Show me sites with high protocol deviations"), removing the technical barrier of SQL.



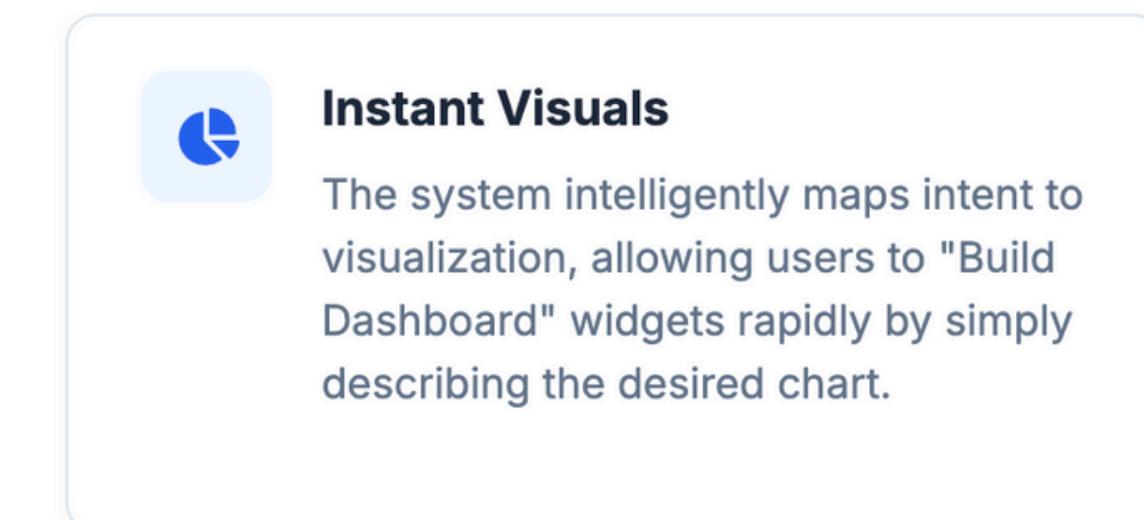
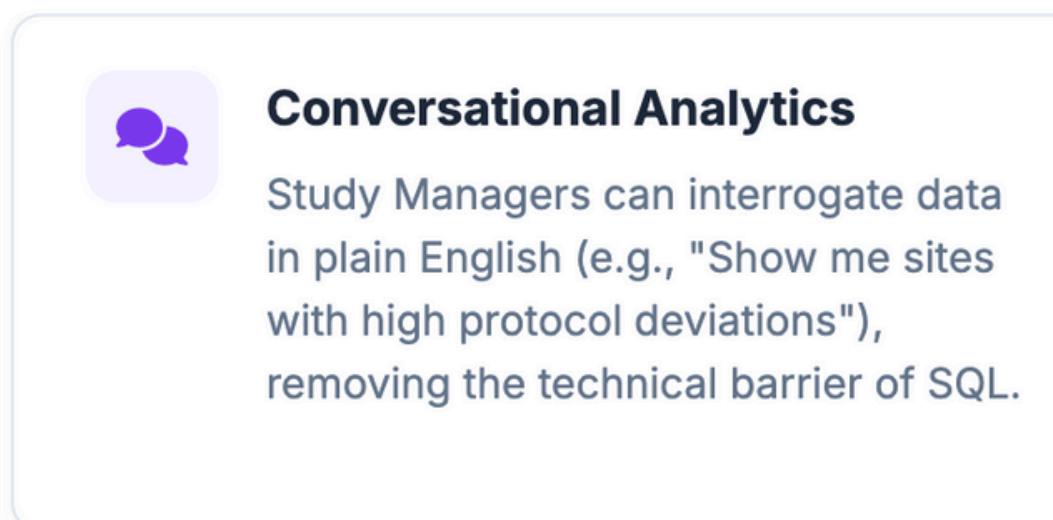
Instant Visuals

The system intelligently maps intent to visualization, allowing users to "Build Dashboard" widgets rapidly by simply describing the desired chart.



Governed Context

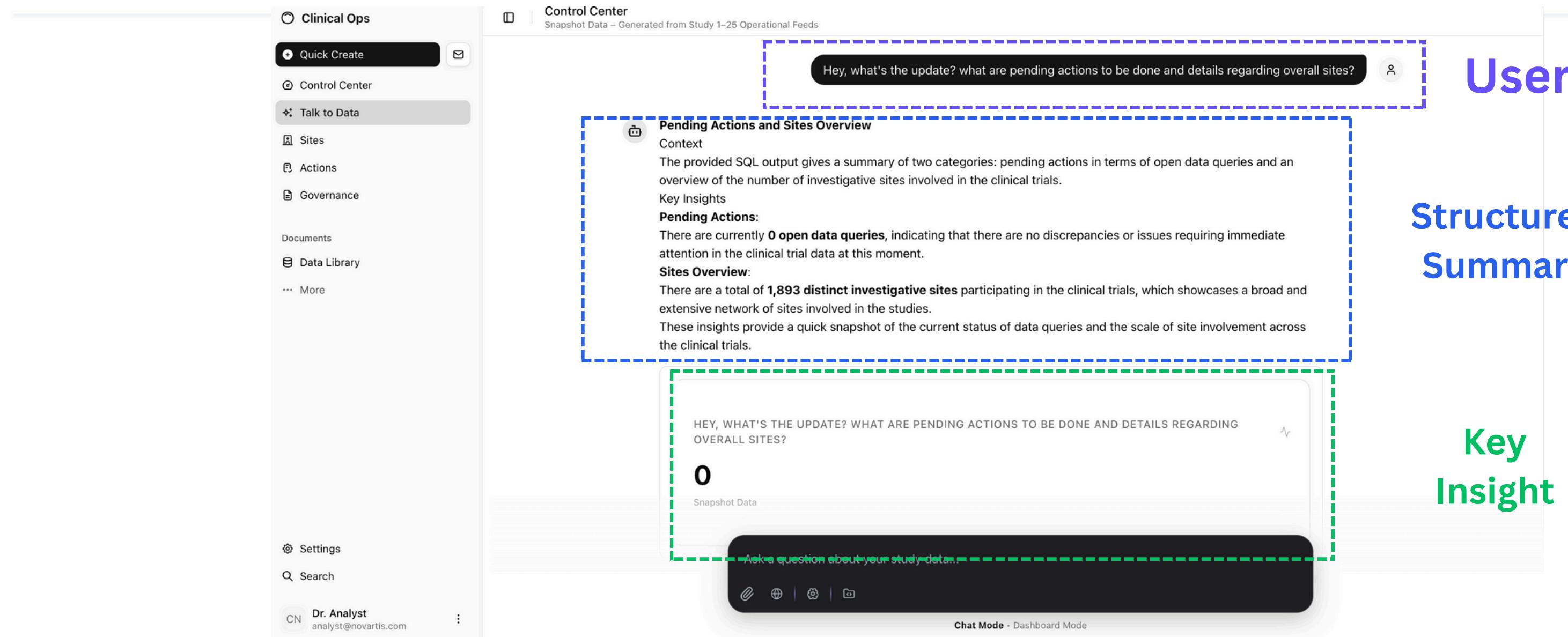
Queries run strictly against the curated "Canonical" data layer. This RAG-based approach prevents hallucinations and ensures answers are inspection-ready.



AI in Action — Grounded Answers

From natural language to structured, actionable insights

 Live Query Execution



The screenshot shows the Novartis Clinical Ops Control Center interface. On the left, there's a sidebar with options like 'Clinical Ops' (selected), 'Quick Create', 'Control Center', 'Talk to Data' (highlighted in blue), 'Sites', 'Actions', 'Governance', 'Documents', 'Data Library', and 'More'. At the bottom left is the user's profile: 'Dr. Analyst' and 'analyst@novartis.com'. The main area is titled 'Control Center' and shows 'Snapshot Data – Generated from Study 1–25 Operational Feeds'. A user query 'Hey, what's the update? what are pending actions to be done and details regarding overall sites?' is entered in a text input field. Below it, a 'Pending Actions and Sites Overview' section provides a summary of pending actions and site participation. A large green dashed box highlights a KPI card showing '0 Open Data Queries'. The interface also includes a 'Chat Mode' button.

User Query

Structured Summary

Key Insight

Intent Recognition

The AI parses the user's complex question ("pending actions... regarding overall sites"), mapping vague terms to specific database entities.

Structured Output

Instead of raw SQL rows, the system returns a narrative summary and specific KPI cards (e.g., "0 Open Data Queries"), providing instant situational awareness.

SQL Translation

Behind the scenes, the NL model generates precise SQL against the "Control Center" snapshot data, ensuring the answer is mathematically accurate.

Safe AI by Design

Controls & Governance

- **Governed Datasets Only:** AI accesses only approved, read-only canonical data; no direct database write access.
- **Strict Isolation:** No training on client data; models are frozen and stateless to ensure privacy.
- **Auditability:** Every prompt, SQL query, and response is logged with user attribution.

Safety Guardrails

- **No Hallucination Risk:** Answers are grounded in retrieved data (RAG); if data is missing, the AI says "I don't know."
- **Evidence Linking:** Every insight includes direct citations to source records or specific data points.
- **Structured Output:** Responses are constrained to specific formats (JSON, standardized text) to prevent drift.

Human-in-the-Loop

AI proposes, but humans decide. The system is designed to augment CRA capabilities, not replace oversight.

- **Review Required:** All critical actions (e.g., site queries, escalations) sit in a draft state until approved.
- **Rejection Feedback:** Human edits improve future signal relevance.

SAFETY STANDARD

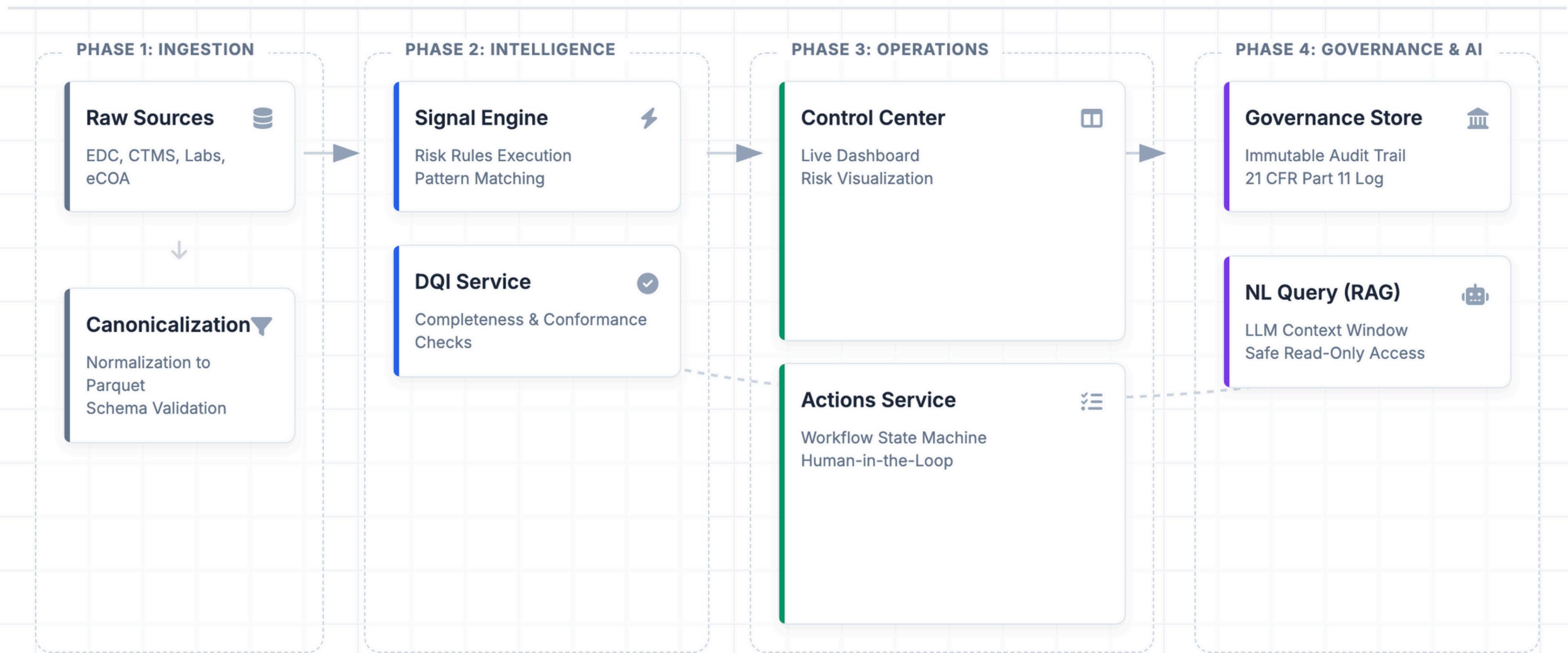
Zero Automated Actions

Without Human Review

Architecture — Under the Hood

Layered operational control system with end-to-end data lineage

 System Topology



● Data Pipeline (Parquet/Arrow)

● Compute Layer (Python/SQL)

● User Application (React)

● AI & Audit Layer

🛡 Zero-Trust Architecture

Value Delivered

Measurable impact across the clinical trial lifecycle

ROI & Outcomes



Operational Efficiency

Faster cycle times, reduced latency

- ✓ **Earlier Detection:** Identify site risks weeks before they become critical issues, preventing study delays.
- ✓ **Reduced Site Escalations:** Proactive intervention reduces the need for "rescue" visits.
- ✓ **Faster Database Lock:** Continuous cleaning means less work at the end of the study.



People & Workforce

Empowering CRAs and Site Managers

- ✓ **CRA Time Saved:** AI drafts initial emails and tasks, reducing administrative burden by ~30%.
- ✓ **Focus on High Value:** Shift from "finding errors" to "coaching sites" and relationship management.
- ✓ **Reduced Burnout:** Automating repetitive triage tasks improves job satisfaction.



Data Quality & Integrity

Higher trust, fewer queries

- ✓ **Higher DQI Scores:** Automated conformance checks drive consistent data entry behavior.
- ✓ **Reduced Missing Forms:** System flags gaps immediately, reducing incomplete subject data.
- ✓ **Lower Protocol Deviations:** Pattern recognition catches systemic misunderstandings at sites early.



Compliance & Governance

Audit-proof and inspection-ready

- ✓ **Inspection Readiness:** Every AI insight and human decision is logged with a timestamp in the Governance Store.
- ✓ **Explainable AI:** No "black box" decisions; regulators can see the exact data lineage for every alert.
- ✓ **Full Traceability:** Link every action back to the specific risk signal and raw data source.

Limitations & Future Scope

Current constraints and the roadmap to production

Strategic Outlook



Current Constraints

Scope of the prototype

- **Snapshot Data Only:** System currently operates on static snapshots rather than real-time streaming ingestion.
- **Pilot Scale:** Validated on a limited subset of subjects (n=1,398) rather than full global trial volume.
- **Synthetic Demo Data:** Demonstrations use anonymized/synthetic datasets to ensure privacy compliance.



Near-Term Roadmap

Production hardening & scaling

- **Production Hardening:** Implementation of robust RBAC, SSO integration, and disaster recovery for enterprise use.
- **Partner Connectors:** Native API integrations for major EDC (Medidata, Veeva) and CTMS platforms.
- **Automated Ingestion:** Fully automated ETL pipelines triggered by source system updates.



Integration Gaps

Areas for connectivity improvement

- **Limited EHR Integration:** Direct Electronic Health Record connectors are not yet implemented.
- **Manual Ingestion Triggers:** Data refresh currently requires manual initiation rather than event-driven webhooks.



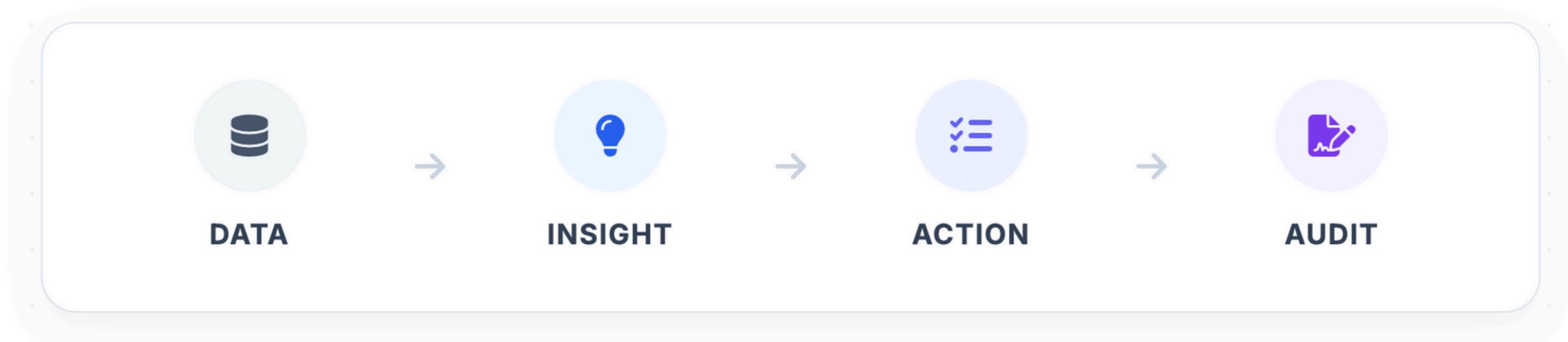
Future Capabilities

Advanced AI & user features

- **Predictive ML:** Transition from descriptive signals to predictive modeling for site risk and data drift.
- **CRA Workload Optimization:** AI-driven assignment balancing based on site complexity and risk profiles.
- **Self-Serve Dashboards:** Natural language "Build Dashboard" capabilities for custom view creation.

More than analytics. A fully **Governed Control System** for Clinical Operations.

We move beyond passive dashboards to deliver an intelligent, closed-loop system that transforms raw trial signals into auditable regulatory evidence.



Thankyou

Team: Pharma Flux



Ready for Deployment

Platform is validated for prospective pilot and regulator-facing dry runs.