TRANSFORMING HEALTHCARE ANALYTICS

With Data Voyager Platform

Powered by Zenera's Self-Coding AI System

Authors

Dr. Maureen Smith, Ravi Kalla (Symphony)
Andrey Ryabov, Ramu Sunkara (Zenera)

Aug 1st 2025

zenera.ai

of healthcare. Built by Symphony and powered by Zenera's Self-Coding AI System, it allows healthcare organizations to rapidly build trusted, curated datasets across EMRs, Claims, HR, and operational systems, without requiring teams of data engineers. One Data Steward, working in plain English, can now achieve in hours what used to take weeks.

2. The Data Bottleneck in Healthcare: Why Clean, Trusted Analytics Remain Out of Reach

Across the U.S., more than **3,500 hospital systems** are racing to modernize their operations using data. But beneath the dashboards and analytics aspirations lies a **foundational problem:** the data itself is fragmented, messy, and painfully hard to work with.

- **Fragmentation of Data Ecosystem:** Each hospital, health system, or specialty group has its own version of the truth, with customized deployments of EMRs like Epic, billing and claims systems, HR platforms, and patient satisfaction tools. Data schemas differ, workflows vary, and fields that should be populated often aren't. Even within the same software platform, implementations diverge by hospital, department, or specialty. What should be a unified source of insight is instead a scattered, inconsistent tangle of systems.

- **Customization Hell:** To extract meaning from this chaos, organizations rely on handcrafted integration pipelines, built by engineers with deep knowledge of APIs, schemas, and EMR internals. These pipelines are brittle, complex, and slow to evolve. A minor change in business logic or reporting requirements can take weeks to implement. Tools like Epic's Clarity and Caboodle offer some structure, but even they require heavy customization and manual effort for each client.

- **Scarcity of Talent:** And the human cost is growing. There's a **shortage of data analysts** who not only understand the technical layers but also the nuances of healthcare systems. Hiring and training these experts is expensive, and keeping them focused on maintenance limits innovation. Meanwhile, analytics teams are stuck waiting for clean, reliable data that rarely arrives on time or in the right shape.

This painful combination of **cost, complexity, accuracy issues, and talent constraints** makes it nearly impossible to scale high-quality analytics across clients. Most healthcare organizations end up working with incomplete data, high turnaround times, and unscalable approaches. The dream of a truly **trusted, curated dataset** remains just that—a dream.

This status quo isn't just costly, it's unsustainable. A new approach is needed, one that redefines how humans and AI collaborate to build healthcare data analytics.

3. Use Cases

We partner with early clients to implement AI-powered workflows that drive measurable outcomes across clinical, operational, and financial domains.

1. **Intelligent Referral Coordination**
    User: Nurse Navigator
    Challenge: Manual intake and follow-up across fragmented systems
    Agentic Workflow:
    - *Automated EHR lookup*
    - *Smart form scheduling*
    - *Patient communication management*

    Impact:
    - *3× faster referral processing*
    - *25% fewer no-shows*

2. **Quality Compliance & Audit Readiness**
    User: Quality/Compliance Lead
    Challenge: Time-consuming audit prep and reporting
    Agentic Workflow:
    - *Extracts structured + unstructured data*
    - *Generates real-time audit summaries*

    Impact:
    - *60% less prep time*
    - *Reduction in manual errors*

3. **Intelligent Information Access**
    User: All Employees
    Challenge: Difficult to find policies and internal knowledge via intranet
    Agentic Workflow:
    - *GenAI maps queries to internal SOPs and policy content*
    - *Delivers contextual, actionable answers*

    Impact:
    - *Saves 5–19 min per employee/day*
    - *Boosts workforce efficiency*

3. Use Cases ...

4. **Value-Based Care Enablement**
    Users: Physicians, Care Teams
    Challenge: Manual coordination of care plans and patient engagement
    Conversational AI Workflow:
    - *Automates patient follow-ups and education*
    - *Integrates EMR, discharge notes, scheduling*
    - *Supports multilingual guidance*

    Impact:
    - *Higher patient satisfaction and outcomes*
    - *Reduced staff burden and payment risk*

5. **Coding & Claims Denial Prevention**
    User: Revenue Cycle Management (RCM) Teams
    Challenge: High billing error rates and denial volumes
    Agentic Workflow:
    - *Predictive coding review*
    - *Pre-claim automation and error correction*

    Impact:
    - *Lower denial rates*
    - *Increased collections and freed-up staff time*

6. **Real-Time Demand–Capacity Matching**
    User: Nursing Staff, Case Managers
    Challenge: LOS and throughput bottlenecks; ED overflow
    Agentic Workflow:
    - *Predicts unit-level demand vs. capacity*
    - *Auto-generates staffing + discharge plans*

    Impact:
    - *More discharges by 2PM*
    - *Lower LOS and fewer LWBS cases*

4. Redefining Data Engineering: One Steward, One Conversation

The Data Voyager Platform introduces a new model: One Data Steward, one conversation with the Zenera AI System. The human works alongside AI, with the AI doing most of the work, replacing a team of experts, data analysts, and developers. The Zenera AI system iteratively builds a model of constraints, validates, and deploys code in a constrained environment.

We take advantage of AI's ability to deliver accurate, context-aware code, often more precisely than human developers. However, humans are still essential in defining domain-specific knowledge and transformation goals. This is not just automation, it's augmentation. AI handles repetitive logic while humans ensure that context, policies, and goals are accurately reflected.

Here's how it works: the human operator interacts with the Data Voyager to iteratively:

1. **Specify input data sources** (e.g., Epic, PeopleSoft, Cerner) and define the format and destination of the curated output.
2. **Explore the structure and metadata** of source systems to identify inconsistencies, missing values, or local customizations.
3. **Design and refine** the extraction, cleansing, and transformation pipelines, guided by AI-generated plans.
4. **Generate and validate** instrumented code, which includes SQL queries, database triggers, parsers, Python, and Shell scripts. This code is equipped with validations, error detection, and runtime observability.
5. **Run and monitor** the system, with real-time feedback loops to improve accuracy and completeness.

The process is dynamic. As the AI system is deployed across different clients, it learns from each environment, recognizing common schemas, recurring issues, and optimal transformation strategies. This accumulated intelligence is reused, accelerating each new deployment.

The result: **One human operator, using only plain English, can build, maintain, and evolve sophisticated data pipelines tailored to each healthcare client.** This turns a costly, slow-moving engineering process into a scalable, conversational workflow, redefining what's possible in healthcare data transformation.

5. A New Way to Build: The AI-Driven Data Pipeline

The figure below and the sample dialogue that follows illustrate how a single human can collaborate with Data Voyager to define input sources, build data pipelines, perform validation, and generate a trusted, curated data store. This workflow is extensible to all hospital systems, regardless of variations in EMR or metadata.

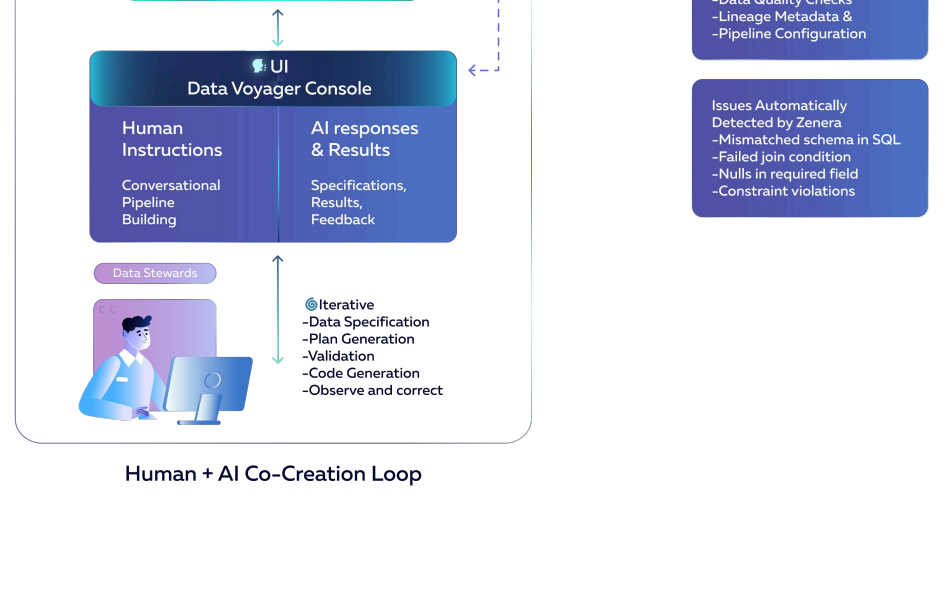

One Human + AI → Clean, Trusted Data at Scale
Fragmented Source Systems (EMR, RCM, Claims, Experience)
Hospital Data Ecosystem (Typical Systems)
Epic
PeopleSoft
Cerner
metadata
Trusted, Curated Data
Data Voyager Data Store
zenera
AI System
(Understand sources, generate & refine plans, produce validated code)
Plan
zenera
Runtime
(Execute & Observe)
Feedback: Runtime Observations → AI + Human Corrections
Zenera Outputs For Every Pipeline
-Transformation SQL Scripts
-Data Quality Checks
-Lineage Metadata &
-Pipeline Configuration
Issues Automatically Detected by Zenera
-Mismatched schema in SQL
-Failed join condition
-Nulls in required field
-Constraint violations
UI
Data Voyager Console
Human Instructions
Conversational Pipeline Building
AI responses & Results
Specifications, Results, Feedback
Data Stewards
Iterative
-Data Specification
-Plan Generation
-Validation
-Code Generation
-Observe and correct
Human + AI Co-Creation Loop

Figure: Zenera AI + Data Voyager transforms healthcare data engineering by enabling a single Data Steward to build, validate, and iterate trusted pipelines via natural conversation and AI guidance—at scale.

Sample Dialogue Between the Data Steward and Zenera AI System

Human Operator (Data Steward)	Zenera AI
Zenera, I need to build a trusted data store for a new hospital client. Their EMR is Epic, and they also have claims data from Optum and patient satisfaction data in Qualtrics. Can you help me set this up?	Absolutely. Let's start by connecting to the data sources. I've identified the APIs and credentials for Epic, Optum, and Qualtrics. Would you like me to crawl the metadata now?
Yes, go ahead.	Done. I've built an initial model of constraints across all three sources. The Epic records have inconsistent date formats, and several fields like "Encounter Type" have missing values. Optum's claim status codes differ from the hospital's internal nomenclature. I've also flagged four columns in Qualtrics that contain local tags that aren't defined in their metadata.
Understood. Let's normalize the date formats to ISO 8601 and fill missing "Encounter Type" fields with a placeholder called "Unknown." For claims, can you map Optum's codes to our standard set?	Mapping complete. I've aligned the Optum claim statuses to the standard ontology. I've also created a transparent mapping document for audit purposes. Would you like to review it?

Human Operator (Data Steward)	Zenera AI
Yes, show me the mapping plan.	Here's the visual and tabular plan. Each transformation is explainable and reversible. Shall I generate the code and set up real-time validation checks?
Yes, generate the code and alert me if more than 5% of any critical column is missing in new data.	Code generated. Pipelines deployed. Observability rules set. I'll monitor the runtime and notify you of any anomalies. Your curated dataset is now available in the Snowflake environment, versioned and ready for downstream analytics.
Perfect. Let's label this pipeline "ClientX_Curated_Store_v1" and prepare a summary for internal review.	Pipeline labeled and summarized. You've just replaced two weeks of manual work with 20 minutes of guided interaction. Shall we proceed to deploy for the next client?

Outcome:

This dialogue illustrates a future where **one person can build what used to take a full team,** using natural language and AI-powered iteration. The Zenera AI System doesn't just automate – it collaborates, explains, adapts, and ensures that the resulting data pipelines are transparent, trusted, and easily maintained.

6. Data Voyager Goals → Components → Workflow

The goals for the Data Voyager are as follows (in the priority order):

1. Build the **trusted data store** for Healthcare clients,

2. Deploy **one use case** (e.g. supply chain optimization or capacity management) and

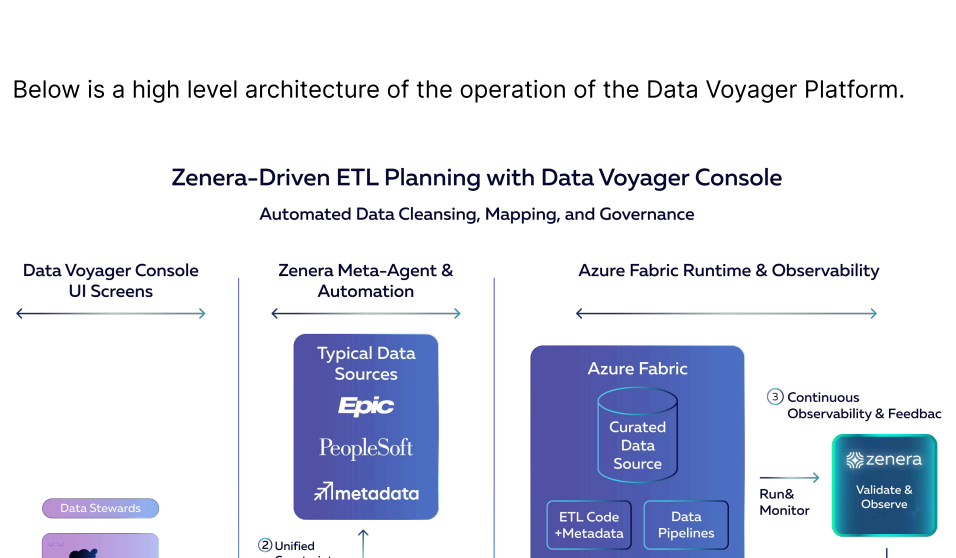
3. Make the curated data **available for third party applications and service providers.**

Building a reliable and accurate Master Patient Index (MPI) is a significant challenge, primarily due to the diversity of data sources, data structures, and terminologies. **Data Voyager** aims to address this complexity by developing **self-learning intelligence, in collaboration with data stewards** from each organization or division. By incorporating a **human-in-the-loop approach,** Zenera will continuously learn and adapt to data variations, including differences in mandatory fields, missing values, and incorrect or inconsistent entries. This collaborative model enables Zenera to build a more accurate understanding of each organization's unique data landscape, ultimately leading to a robust and trustworthy MPI.

Data Voyager Components:

- **Data Voyager Console** – A user-facing interface for Symphony Data Stewards to configure pipelines using natural language.
- **Zenera Meta Agent** – Powers Data Voyager's self-coding engine by interpreting inputs, crawling APIs, generating plans, and producing validated code.
- **Observability & Validation Layer** – Monitors data pipelines for accuracy, errors, and schema changes, enabling real-time refinement with human intervention.

Below is a high level architecture of the operation of the Data Voyager Platform.


Zenera-Driven ETL Planning with Data Voyager Console
Automated Data Cleansing, Mapping, and Governance
Data Voyager Console UI Screens
Zenera Meta-Agent & Automation
Azure Fabric Runtime & Observability
Typical Data Sources
Epic
PeopleSoft
metadata
Azure Fabric
Curated Data Source
ETL Code +Metadata
Data Pipelines
③ Continuous Observability & Feedbac
zenera
Validate & Observe
Run& Monitor
Data Stewards
② Unified Constraint Modeling
zenera
Meta Agent (Builds Constraints, Plans & Generates ETL Code Iteratively)
Deploy SQL, Data Quality Checks, Metadata
UI Step ① Input Data, Metadata, Output
Specify e.g., EMR data, HR data
UI Data Voyager Console
① Conversational Specification
UI Step ② Constraint Modeling e.g., Single Code Space
UI Step ③ Plan Generation & Feedback cycle
Plan e.g., code for null checks
③ AI-Driven Planning and Validation
④ Code Generation
UI Step ④ Code Generation & Validation
Generate Code e.g., create SQL joins
Generated Code Artifacts:
-Transformation SQL Scripts
-Data Quality Checks
-Lineage Metadata (Purview) &
-ADF/Synapse Pipeline Configuration
UI Step ⑤ Validate & Observe e.g., detect schema mismatches
Errors
-Mismatched schema in SQL
-Failed join condition
-Nulls in required field
-Constraint violations
-Errors
-User-approved
-AI-inferred
-Feedback loop
-Runtime error detected & fixed

Workflow: Five-Step Self-Coding Pipeline Development

1. **Conversational Specification** – Data stewards use plain English to define data sources, transformation goals, and target destinations.
2. **Constraint Modeling** – Zenera Meta Agent ingests documentation, APIs, and metadata to build a contextual model.
3. **AI-Driven Planning –** Suggested plans for transformations and analytics are reviewed and adjusted by users.
4. **Code Generation** – Zenera writes pipeline code (SQL, scripts, ETL jobs) that is production-ready. Zenera offers testing capabilities in the studio to validate the results.
5. **Observability Feedback Loop –** Ongoing validation identifies data issues and supports rapid reconfiguration. In addition, Zenera logs all the execution results for auditability.

6. How Zenera Powers the Data Voyager Stack

Each layer of the Data Voyager Stack—Enterprise Data Core, Knowledge Layer, Agentic Automation, Experience Layer, Governance—is mapped to a core capability in Zenera's architecture.

Data Voyager Stack	Zenera Architecture Mapping (Terminology & Description)
Enterprise Data Core	**Corpora** – provides definitions for both static and dynamic data sources. It automatically scans static documents, uses APIs, and generates SQL code to access dynamic data, thereby building the Enterprise Data Core. Data Stewards can define the Enterprise Data Core via a conversational specification (this is the first step in the Data Voyager Console). Zenera then builds a Model of Constraints for each client's data core (including a Single Code Space for structured data and a Semantic Database for semi-structured and unstructured data).
Knowledge Layer	**Transforms** – allow Data Stewards (humans) to guide the AI model's understanding of the data. These transforms tell the system how to process and interpret data – for example, how to match user queries to the underlying data and how to map or use metadata. Transforms can be specific to each client's needs; Data Stewards define the ontology and synonyms relevant to the data and its metadata for their organization.
Agentic Automation Layer	**Reasoning Graphs** – represent the AI-driven orchestration of data retrieval and processing. The self-coding system automatically decides which corpus to query, how to match and transform the data, and it encodes these decisions in a reasoning graph. If necessary, a Data Steward can modify these AI-generated reasoning graphs using natural language instructions to refine the outcomes.

Data Voyager Stack	Zenera Architecture Mapping (Terminology & Description)
Experience Layer	**Meta Agent** – determines how to present data and insights to end users through the user interface. For example, the Meta Agent may choose to display results as a chart, graph, table, or form depending on the user's role and context (a Transitional Care Nurse might see a patient dataset as a chart or form, for instance). These presentation formats (UX) are not fixed – Data Stewards can alter how information is shown by adjusting the reasoning graphs, thus influencing the Meta Agent's behavior.
Governance Layer	**User Proxy & Governance** – at runtime, Zenera acts as a proxy on behalf of the user, enforcing security and access controls. It will only retrieve or manipulate data if the user has the appropriate permissions to do so. Moreover, additional governance rules (for example, policies from an external system like Microsoft Purview) can be integrated into the reasoning graphs, ensuring that all data access and operations comply with the organization's governance requirements.

7. Economic and Strategic Impact

- A team of 10 analysts can manage 10+ custom deployments in parallel.
- Weeks-long development timelines shrink to hours or days.
- Reduces dependence on scarce engineering talent.
- Higher-quality, cleaner data leads to better decisions and outcomes.
- Scalable across departments, hospital systems, and health networks.

Data Voyager enables 10x faster data pipeline deployment, cuts operational costs, and eliminates backlogs by reducing team size. It will scale to hundreds of hospital clients without sacrificing control or accuracy.

8. Potential Business Model

- **TAM:** 3,500+ U.S. hospitals.
- **Fixed License Fee:** Annual Data Voyager subscription starting at $1M+.
- **Variable Revenue:** Shared savings model tied to realized cost reductions (e.g. in supply chain, operational efficiencies). In addition, as the AI infrastructure costs come down and also reduces maintenance manpower as the system is almost 100 % reliable we can pass part of these benefits to customers.

The platform supports enterprise licensing or usage-based pricing. Deployment can start with a single system, with measurable ROI at each phase.

9. Open Ecosystem Architecture

- Data Voyager ships with starter use cases (e.g. supply chain, capacity management).
- Allows clients and third parties to build their own apps on top of Data Voyager curated datasets.
- Compatible with cloud and on-prem compute environments (e.g., Azure).
- Ensures governance, privacy, and explainability out-of-the-box.

Unlike closed vertical stacks, Data Voyager empowers Symphony and its clients to build on top of curated data with full control

10. Comparison: Data Voyager vs. Palantir Hospital 360

Both aim to transform healthcare analytics. Only one is built for scalable, human-AI collaboration.

Feature	Data Voyager	Palantir Hospital 360 (AIP)
Human Operators Required	**One Data** Steward using a conversational interface.	**Many** – larger teams are likely required for deployment.

Feature	Data Voyager	Palantir Hospital 360 (AIP)
Data Connectivity	**Zenera "Corpora"** enables automated connection to both static and dynamic datasets. Generates access code programmatically, supporting a wide range of systems (EMR, ERP, CRM, etc.) and allowing fast harmonization of data.	Offers **specialized data connectors** for integration with known systems.
Time to Deploy	**10x faster** via Zenera's **self-coding AI.** Deployments follow a five-step conversational process with the Meta Agent, making it highly agile and iterative.	Slower, with **manual configuration** processes. Specific deployment timelines are unclear.
Custom Ontology Support	**Yes.** Data Stewards can define and update ontologies directly using natural language through "Transforms".	**Yes, but** ontologies must be **hard-coded,** requiring more effort and technical expertise.
Tailored to Each Health System	**Automatically builds a constraint model** specific to each healthcare client. The AI then generates customized code for extraction, mapping, and transformation to produce a curated dataset.	**Manual integrations** required for each new deployment, including data source mapping and customization.
AI Explainability & Auditability	**Built-in.** Reasoning Graphs and Transforms ensure transparency, traceability, and auditability of AI-generated logic and outputs.	**Unknown.** Explainability features are not publicly detailed.
Ecosystem Openness	**Open and extensible.** Partners can license the Data Voyager platform to build and embed their own solutions on top of the curated data foundation.	**Closed ecosystem.** Palantir offers vertically integrated solutions developed in-house.

11. Conclusion and Call to Action

Symphony's vision to deliver a next-gen healthcare analytics solution becomes real with Data Voyager, powered by Zenera's AI Platform. This strategic stack delivers for each Healthcare Client:

- A scalable, domain-specific analytics system.
- Unprecedented speed and cost-efficiency.
- Human-in-the-loop governance and agility.

We propose a follow-up meeting to explore a deployment of a co-branded solution that will unlock value for Symphony's hospital clients and position Symphony as a leader in intelligent healthcare transformation.

Let's modernize healthcare analytics—starting today.