Paper SD02

Dismantling the Black Box:

Why We Need to Think Bigger About AI Integration

Christina Dinger, ThoughtSphere, San Jose, United States

Joe Lebers, ThoughtSphere, San Jose, United States

ABSTRACT

AI is transforming the clinical development value chain by generating insights, reducing silos, and automating clinical data transformation and review tasks. However, when AI models are deployed without seamless integration, traceability, and explainability, the value they bring is often lost. The key to effective AI is when it functions as an intuitive, seamless extension of existing workflows.

This paper explores how AI can become a "new stakeholder" in the clinical trial process. It will have its own vested interest, priority, relationships, trust, and value, but will contribute to the collective team, not be isolated in a black box.

We'll demonstrate how ThoughtSphere's platform enables a collaborative relationship between humans and machines, delivering advanced data aggregation and transformation, real-time analytics, and automated data check generation. Our integrated AI/ML approach minimizes hallucinations, eliminates redundancies, and streamlines clinical data operations from study build to submission, ensuring a unified, efficient process.

INTRODUCTION

"Wherever the art of medicine is loved, there is also a love of humanity", Hippocrates. Medicine—and by extension, clinical development—exists at the intersection of science and art. Science provides the foundation, built on quantifiable data, established methodologies, and historical precedent. But the art of medicine lies in human intuition, empathy, the ability to envision possibilities beyond the known, and the creative problem-solving required to navigate uncharted medical frontiers.

AI's role in clinical development must respect this duality. The true power of AI in clinical trials is unlocked when it works collaboratively with human expertise, combining data-driven insights with human imagination, ethical reasoning, compassion, and strategic thinking in a seamless, collaborative manner – rapidly extrapolating key scientific insights to fuel future innovation.

However, AI use in clinical trials today remains largely siloed and task specific. The models lack the context, adaptability, and interactivity necessary for true transformation. For AI to truly revolutionize clinical trials, it must go beyond isolated functionalities. It must function as a unified and dependable collaborator that integrates seamlessly into the clinical trial ecosystem. This paper delves into the current state of AI in clinical trials, explores its trajectory, and demonstrates how ThoughtSphere's platform redefines AI integration to offer a transparent and efficient AI-driven experience.

THE STATE OF AI IN CLINICAL TRIALS TODAY

Digital transformation and AI-driven innovations have become pivotal to modern clinical development. The global AI market in clinical trials is projected to expand from $1.42 billion in 2023 to $8.5 billion by 2035, reflecting a compound annual growth rate (CAGR) of 16% (Healthhark, 2024). Despite this significant investment and the industry's eagerness to revolutionize clinical trials, most AI models in use today remain focused on isolated, single-use cases within the clinical development value chain—such as protocol design, patient recruitment, anomaly detection, and safety signal identification.

While these single use case models can (and do) deliver efficiencies and new insights, the "spotty" and fragmented AI approach adds further complexity to the very fluid and intricate clinical development process -- where one variable or change can have a cascade effect on other events and variables. The efficiency gains of single use case models are often hampered by scalability challenges, limited model scope, and the generation of false positives. This leads to skepticism about AI's true value and perhaps is a reason behind the slower pace of innovation according to a 2024 digital disruption survey (ICON, 2024). The limitations of today's AI in clinical trials can generally fall into one of three key challenges:

1. PROCESS FRAGMENTATION & REDUNDANCY

Many AI models function in parallel to existing operational systems and workflows rather than being embedded within them. This fragmentation leads to:

- Duplicative Efforts – AI models flag the same discrepancies already identified by project teams, wasting time and resources.
- False Signals & Alert Fatigue – AI models working in isolation lack sufficient data context, increasing the risk of false positives. End-users are forced to sift through irrelevant alerts, reducing efficiency rather than improving it.
- Limited Operational Impact – Because AI outputs often exist in separate dashboards or reports and were trained on limited data scopes for a particular purpose, they require manual interpretation and action from humans who see the broader data landscape.

2. SCALABILITY ISSUES & TRAINING BURDENS

AI models in clinical trials often require extensive upfront training using historical clinical and real-world evidence (RWE) data. This presents several scalability challenges:

- High Resource Demand – Model training requires significant processing power, time, and curated datasets, making it difficult to scale across multiple trials.
- Poor Suitability for Early-Phase Studies – First-in-human (FIH) and rare disease trials suffer from insufficient training data, leading to AI hallucinations or models that are simply not deployable due to lack of training data.
- Limited Adaptability – Even after training, AI models often struggle to generalize across different trial designs, therapeutic areas, or regulatory requirements—further limiting scalability.

3. LACK OF EXPLAINABILITY & TRUST ISSUES

One of the biggest challenges with AI in clinical trials is a lack of transparency in how AI generates findings. Without clear explanations, even clinically valid AI-generated insights may be dismissed by trial teams.

- Opaque Decision-Making – Many AI models fail to provide traceability into how they reached a conclusion, making it difficult for stakeholders to trust or act on recommendations.
- Missed Opportunities – Even when AI uncovers statistically relevant correlations that may go unnoticed by clinicians, if it doesn't provide justification, the project team may disregard the finding as a false positive or hallucination. For example, if an AI model surfaces a potential drug-safety signal, but does not explain why or provide supporting evidence, medical monitors may disregard the finding due to lack of confidence—potentially missing a critical safety risk.
- HITL Dependency Without Clarity – Human-in-the-loop (HITL) feedback mechanisms are critical for AI learning, but when models work in isolation end-users spend more time sifting through false-positives and educating the AI model than benefiting from its insights.

THE PATH FORWARD: FROM SILOED AI TO INTEGRATED INTELLIGENCE

For AI to truly transform clinical trials, it must evolve beyond fragmented, single-use applications into an integrated, contextual, and explainable intelligence layer that supports decision-making across the trial continuum. *"Strategic combinations of digital tools can result in solutions that are greater than the sum of their parts. This is especially true for end-to-end solutions that bridge study stages to ensure smoother startup, operations and closeout to maximize the study timelines, data quality and ultimate ROI."* (ICON, 2024). For AI to evolve to this state, a strong data foundation to support operational, clinical, and RWE data across the clinical development life cycle is required.

BUILDING THE RIGHT ECOSYSTEM FOR AI EVOLUTION

The foundation for unlocking AI's full potential lies in democratizing data—making it accessible, understandable, and actionable for a wide range of users and use cases. A well-designed data fabric enables seamless data exchange and real-time access to both structured and unstructured data, empowering end-users and AI models alike. Yet only 20% of organizations have an architecture robust enough to support AI workloads (Economist Impact, 2024). The first step to deploying successful AI models is establishing a scalable, cloud-native data architecture capable of ingesting, normalizing, and storing heterogeneous data types and formats. This architecture must support dynamic resource allocation and parallel processing to handle high-velocity data streams and evolving computational demands.

Layering the rich data fabric with end-to-end data orchestration processes and workflows creates a seamless interface for data transformation, interpretation, cleaning, and analysis dataset generation. It fosters collaboration across end-users and facilitates a holistic, Quality-by-Design approach. It also creates an ideal environment for AI models to learn and interact with data and users across the clinical development life cycle to develop deep neural networks. This is where enhanced AI integration becomes possible.

Central methods to support AI's evolution are:

- Embedding AI within operational workflows – AI should not function in parallel to human-led processes but be deeply integrated into clinical workflows, automating steps when possible. "Just-in-time" HITL quality gates should be enabled to safeguard AI-driven outputs throughout the process. For example, when transforming raw data to SDTM, the model automates variables and lets the end-user know the confidence that the variable was mapped accurately and provides an intuitive workflow for the human to verify the mappings before the SDTM data sets are approved.

- Enabling continuous, cross-functional learning – Traditional AI models rely on static, upfront training datasets, making them rigid and prone to failure in new clinical scenarios. Instead, AI should dynamically learn from real-time trial data, how that data is transformed and used, investigator feedback, and regulatory changes, ensuring its intelligence evolves in sync with the clinical landscape. This rapid learning is only possible when the model is exposed to cross-functional interactions and data use cases.

- Ensuring Explainability and Transparency – Trust in AI hinges on explainability (XAI). Built-in explainability trees and model confidence scoring allows users to understand the rationale behind AI-generated recommendations and how confident the model is about the output generated. By incorporating confidence intervals and traceable reasoning, AI can provide clinicians, data managers, and statisticians with meaningful, actionable insights rather than opaque predictions.

ESTABLISHING SAFEGUARDS TO SUPPORT MODEL MATURATION

It is only through contextual learning and continuous maturation that AI can truly evolve. This requires putting up guard rails and quality mechanisms to support AI models as they mature. Below are three well-respected approaches to enhance AI's ability to provide relevant, reliable, and continuously improving insights in a highly regulated domain like clinical research.

Retrieval-Augmented Generation (RAG): Traditional AI models generate responses based purely on their pre-trained knowledge, often leading to hallucinations or outdated insights. Retrieval-Augmented Generation (RAG) enhances contextual relevance by dynamically pulling real-time, context-specific data from external knowledge sources before generating a response. This is like passing the data through an additional lens, so the model understands the data in the proper context. In clinical trials, RAG allows AI to:

- Reference up-to-date regulatory guidelines, coding dictionaries, lab normal ranges, trial protocols, and patient data instead of relying solely on static training data.

- Adapt responses based on study-specific requirements, ensuring trial data is interpreted within the right clinical and statistical context.

- Reduce hallucinations by grounding AI outputs in authoritative, verifiable sources.

By integrating RAG, AI becomes a more reliable decision-support tool, able to generate insights that reflect the latest available evidence rather than outdated or generalized information.

Just in Time HITL: Despite AI's advancements, human expertise remains irreplaceable, especially in clinical trials where new drug discoveries and unpredictability of patient outcomes converge at the art of medicine. HITL mechanisms ensure that machine-driven insights are continuously validated, refined, and sometimes refuted by expert oversight through the form of "quality gates" in the model's processing journey. While AI and LLMs are able to process vast amounts of data quickly to uncover patterns that humans may miss, they lack the ability to generate "new knowledge" which is ultimately the objective of clinical research. *"Human cognition is driven by forward-looking, theory-based causal logic, which is distinct from the emphasis AI and computational models of cognition place on prediction and backward-looking data*." (Felin, Holweg, 2024). Thus, AI models working autonomously or in a vacuum, separate from humans is not conducive to advancing clinical development. HITL must be incorporated throughout the process chain to provide just-in-time:

- Quality control and error correction—Clinical data managers, statisticians, and medical monitors can review and adjust AI outputs *before* they influence critical decisions.

- Ethical and regulatory compliance—HITL ensures that AI-driven recommendations adhere to Good Clinical Practice (GCP), ICH guidelines, and FDA/EMA requirements.
- Adaptive learning—By integrating expert feedback, AI systems progressively improve their accuracy and relevance, learning from human corrections in real-time.

While reliance on HITL validation may decrease as AI becomes more robust, it remains essential during the early stages of AI adoption (ACRP, 2025). Additionally, it provides a layer of traceable accountability in the process.

Bayesian Inference: Clinical trials operate in a dynamic landscape, where new data constantly emerges from patient responses, adverse events, and evolving trial parameters. *"Bayesian and related approaches to cognition emphasize the dynamic updating of beliefs by which prior knowledge (a prior) is integrated with new evidence to revise beliefs (posterior) in a process mathematically described by the Bayesian formula (Pinker 2021). This iterative updating, reflecting a continual learning process, acknowledges and quantifies uncertainty, framing understanding and decision making as inherently probabilistic."* (Felin, Holweg, 2024). Unlike static AI models, Bayesian Inference enables AI to continuously update its predictions as new data becomes available. This approach:

- Uses probabilistic reasoning to refine AI predictions based on the most recent evidence.
- Adjusts AI-generated insights in real time, allowing for adaptive risk assessments, early signal detection, and data-driven trial modifications.
- Supports personalized clinical insights, as Bayesian models refine probability estimates based on evolving patient characteristics and treatment responses.

In clinical data management, Bayesian methods strengthen AI's ability to handle uncertainty, making it more adept at recognizing emerging patterns, adjusting risk models, and guiding decision-making with up-to-the-minute accuracy.

By fostering a dynamic and interactive AI ecosystem, efficiencies will emerge naturally, and AI models will progress from passive tools to active, adaptive intelligence. This paves the way for the next leap forward in AI's role in clinical trials—the rise of Agentic AI.

THE RISE OF AGENTIC AI: INTERACTIVE, CONTEXT-AWARE, & PROACTIVE

The future of AI in clinical trials isn't just about making AI smarter—it's about fostering a true partnership between AI and human experts. Agentic AI, which refers to artificial agents with natural language interfaces capable of planning and executing sequences of actions on behalf of users, is poised to become a central feature of daily life. Unlike current text-to-speech applications like Alexa and Siri, these agents will be far more sophisticated, providing actionable insights and executing complex tasks in clinical trial settings (Economist Impact, 2024).

In this evolving paradigm, AI will no longer act as a passive tool but as an active member of the clinical team, engaging in two-way interactions with clinicians, statisticians, and data managers. *"Firms have an ability to organize multiple parties with diverse expertise, incentives, and knowledge to work toward shared objectives. A recent addition to this collaborative ecosystem is the digital employee, or AI agent. AI agency has the potential to fundamentally transform the structure and functioning of firms."* (Canayaz, 2025) Through interactive dialogue and iterative learning, AI agents will help shape strategies, offering more than just surface-level insights. This collaborative approach will bridge clinical, operational, and statistical workflows, uncovering hidden patterns across the clinical development value chain to enhance team efficiency, surface dynamic risks, and address safety concerns more swiftly.

Crucially, AI must be embedded across all clinical data processes—from data ingestion to real-time analysis—forming a deeply interconnected system rather than existing as fragmented tools. This end-to-end integration will allow AI to move beyond identifying correlations and begin recognizing true cause-and-effect relationships. Context awareness will be key, as AI continuously learns from trial data, regulatory requirements, and investigator feedback, refining its recommendations over time to become increasingly precise, proactive, and adaptable to new complexities.

Ultimately, AI will evolve from a passive automation and analytics tool into a trusted partner that grows alongside clinical teams. With the right infrastructure—such as ThoughtSphere's end-to-end data orchestration platform—this transformation from reactive automation to proactive collaboration is not just possible but inevitable.

THOUGHTSPHERE: THE IDEAL ECOSYSTEM FOR AI EVOLUTION IN CLINICAL TRIALS

AI's true potential in clinical trials lies not just in automation but in seamless integration and interactive collaboration. The ThoughtSphere Data Orchestration Platform offers the perfect environment for AI to evolve into a proactive, decision-supporting team member across clinical development.

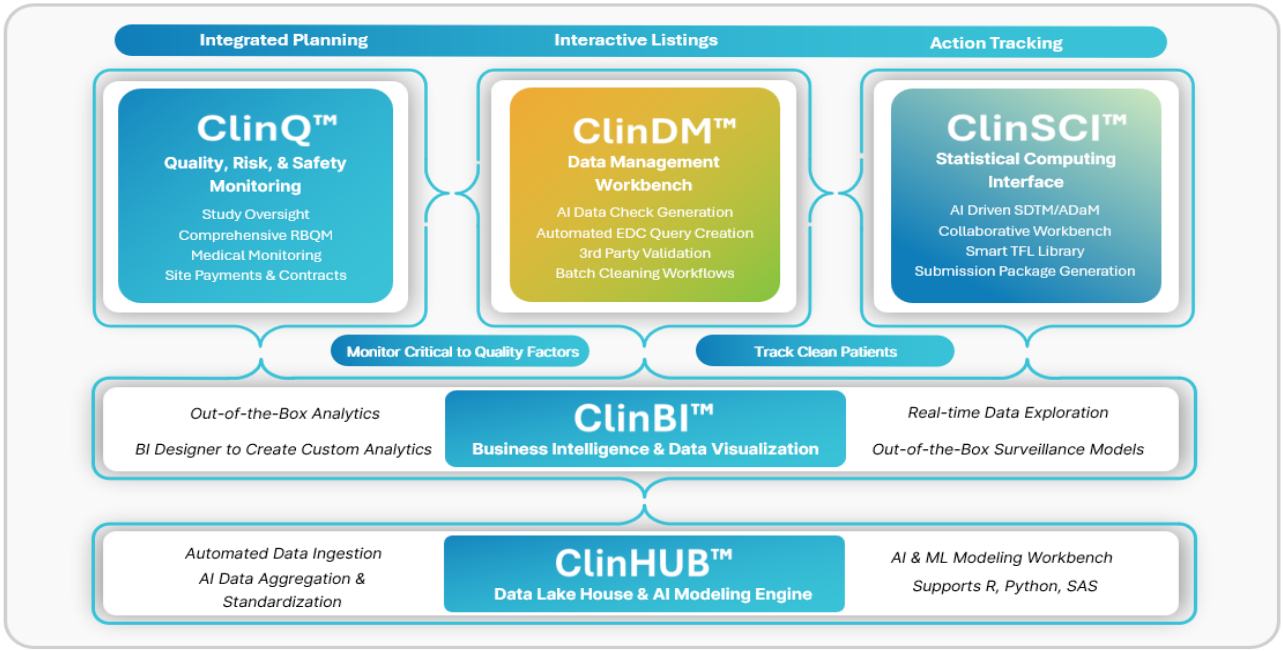
A UNIFIED DATA PLATFORM FOR AI-DRIVEN INTELLIGENCE

Just as the human brain synthesizes multiple sensory inputs for comprehensive understanding, AI in clinical development must function within a centralized, connected ecosystem. However, most organizations are far from achieving this level of integration—recent research reveals that only 22% of organizations have a data architecture capable of supporting AI workloads without modification, while 48% of data engineers spend the majority of their time resolving data source connections (Economist Impact, 2024).

ThoughtSphere's stacked platform was built to solve this problem. ClinHUB™, its AI-driven data lake and modeling engine, centralizes data processing and supports seamless data ingestion, standardization, and real-time AI-driven insights. This cloud-agnostic, scalable infrastructure eliminates silos and creates an ideal foundation for end-users and AI to collaborate and continuously learn through integrated workflows.

THE POWER OF INTERCONNECTED AI ACROSS THOUGHTSPHERE MODULES

- Centralized Data Processing (**ClinHUB™**): Acts as the data backbone, integrating data from multiple sources and enabling real-time transformation, AI modeling, and predictive analytics.
- Business Intelligence & Visualization (**ClinBI™**): Delivers no-code dashboards with 150+ built-in reports, helping users convert AI-driven insights into actionable decisions across studies.
- Integrated Risk Management & Oversight (**ClinQ™**): Provides real-time quality and safety monitoring with AI-assisted insights for proactive risk management and automated tracking of site payments and contracts.
- AI-Enhanced Data Management (**ClinDM™**): Automates data cleaning, EDC query generation, and compliance workflows, ensuring AI works with clean, standardized datasets.
- Collaborative Statistical Analysis (**ClinSCI™**): Offers an intuitive statistical computing interface with pre-built templates for SDTM/ADaM datasets and TLF generation using R and Python™, empowering users to validate AI-assisted outputs.



DEEP CONTEXTUAL LEARNING & NEURAL NETWORK EVOLUTION

ThoughtSphere's end-to-end platform ensures that AI continuously learns across the entire clinical data lifecycle. Every data ingestion, transformation, and user decision contribute to refining the platform's data ontology and the AI model's contextual understanding and predictive accuracy. This interconnected ecosystem allows AI to evolve from a static tool into an intelligent system that adapts to trial complexity, recognizes patterns, and delivers increasingly precise insights.

DYNAMIC WORKFLOWS WITH HITL VALIDATION CHECK POINTS

ThoughtSphere's integrated platform embeds AI into every stage of the clinical trial process—from initial data collection to regulatory submission. HITL validation checkpoints are strategically placed at key junctures during study planning and start-up workflows. These checkpoints allow end-users to review and verify AI-generated recommendations before they inform critical outputs such as datasets, study risk assessments, and validation reviews. This collaborative approach builds user confidence, ensures accuracy, and continuously refines the model's performance through an iterative improvement loop.

THE FUTURE OF AI IN CLINICAL TRAILS: THOUGHTSPHERE'S VISION

Emerging trends in AI—such as regulatory-ready AI, unified data platforms, and interactive AI agents—are shaping the next generation of clinical trial intelligence. ThoughtSphere's data orchestration platform is uniquely positioned to lead this evolution by embedding AI into end-to-end workflows, allowing AI to learn dynamically, collaborate with clinical teams, and support data-driven decision-making.

By transforming AI from a fragmented, single-use tool into an integrated intelligence layer, ThoughtSphere is redefining what AI can achieve in clinical development, bridging the gap between automation and human expertise to drive the next era of innovation in medicine.

REFERENCES

Association of Clinical Research Professionals (ACRP). (2025, January). *Responsible Oversight of Artificial Intelligence for Clinical Research Professionals* Responsible-Oversight-of-AI-ACRP.pdf

Economist Impact. (2024, November). *Unlocking enterprise AI: Opportunities and strategies* (Sponsored by Databricks). Economist Group. A global study of 1,100 technologists—plus 28 CIO interviews | Databricks

Healthhark. (2024, July). *The Transformative Role of Artificial Intelligence (AI) in Clinical Trials* https://healtharkinsights.com/the-transformative-role-of-artificial-intelligence-ai-in-clinical-trials/

ICON. (2024, December). *Digital disruption: Surveying the industry's evolving landscape*. ICON plc. Digital Disruption: Surveying the industry's evolving landscape

Mehmet Canayaz (2025, January). *AI Agency.* Pennsylvania State University - Smeal College of Business https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5109326

Teppo Felin, Matthias Holweg (2024). *Theory Is All You Need: AI, Human Cognition, and Causal Reasoning*. Strategy Science 9(4):346-371. https://doi.org/10.1287/stsc.2024.0189

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Author Name: Christina Dinger

Company: ThoughtSphere

Address: 99 S Almaden Blvd, San Jose, CA 95113

Email: christina.dinger@thoughtsphere.com

Website: thoughtsphere.com

Author Name: Joe Lebers

Company: ThoughtSphere

Address: 99 S Almaden Blvd, San Jose, CA 95113

Email: joe.lebers@thoughtsphere.com

Website: thoughtsphere.com