AI Agents for Accelerating Data Analysis in Clinical Development

PharmaSUG SDE – Boston, Oct 15th 2025

Weijie Zhang, Ph.D., Zoe Wu, and Xing Chen, Sc.D.

moderna

The Way We work

People

- Statisticians , Data Scientist , Data Engineer
- Builder mindset
- Cross-functional domain experts

Process

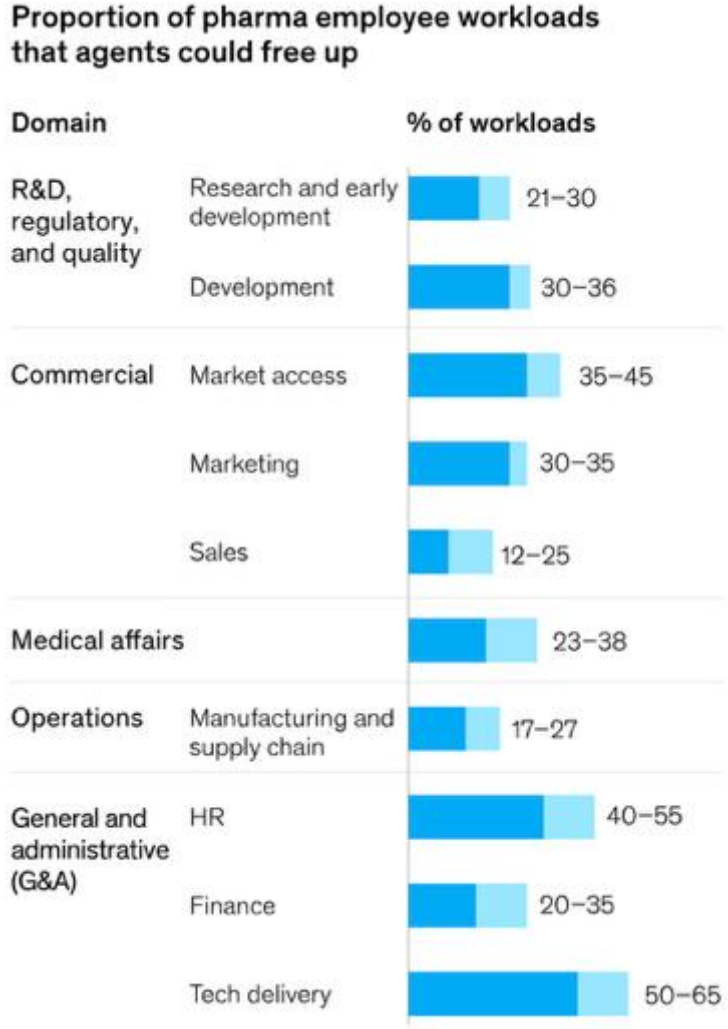
- TLF and cross study data needs
- Ongoing statistical needs & agile workstreams
- Ad-hoc exploratory analyses
- Translational analyses

Technology

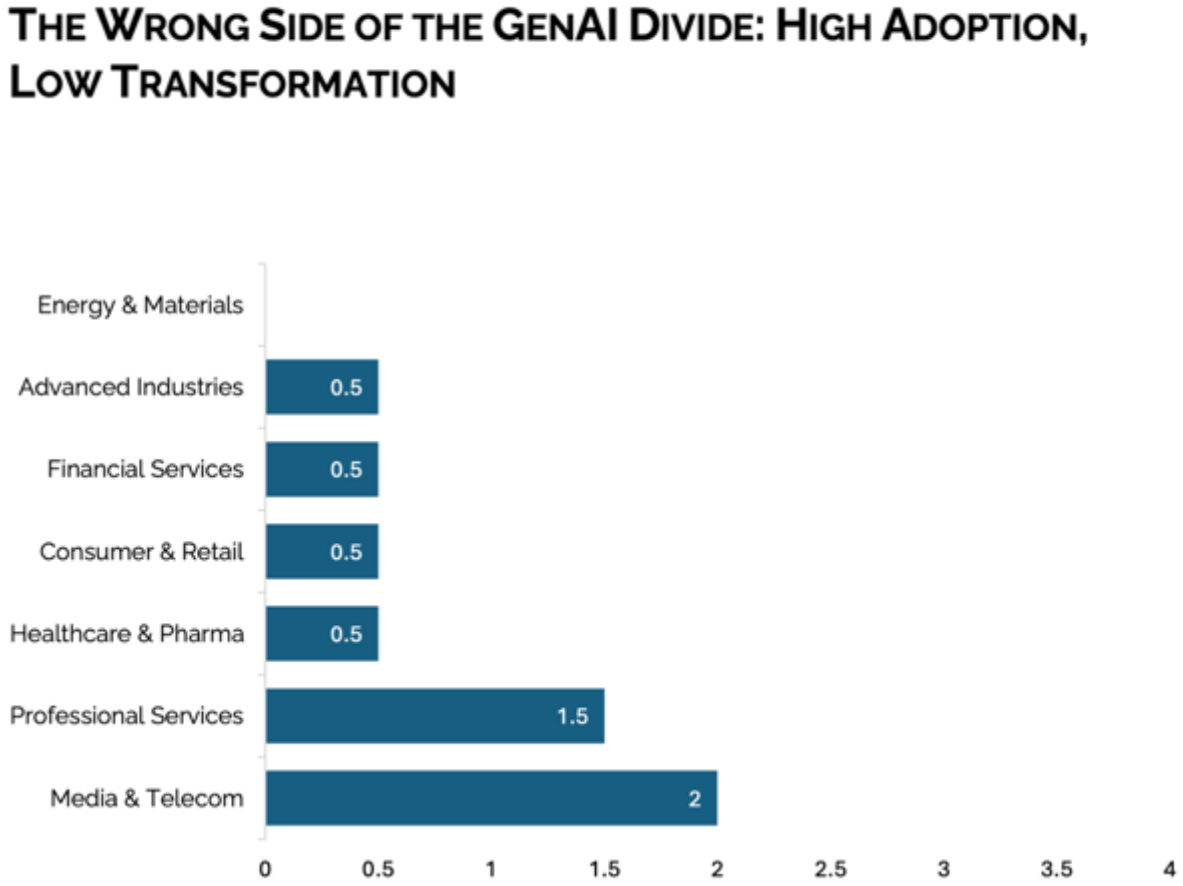
- R/Python
- ML/GenAI
- Cloud, Database
- APIs and Pipeline integration and automations

Agentic AI reshapes pharmaceutical industry

Potential benefits from AI agents:

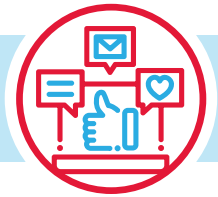


Data from: McKinsey & Company



Adopted from MIT NANDA: STATE of AI in Business 2025

The AI Playbook–three major scenes of AI use in industry



Chatbots

Conversational Q&A applications. ChatGPT, Claude, Gemini, etc.

- Has become a staple item in everyday work.
- Suits diverse business settings and functions.
- Speeds up information sharing for specific tasks.

Copilots

Assistive tools embedded into workflows (e.g., Github Copilot, MS Copilot)

- Handles repetitive tasks.
- Helpful tool for programmers, statisticians, computational scientists, etc.

Agents

Autonomous systems that can plan, call APIs, and execute workflows end-to-end.

- Requires deep understanding of LLM frameworks.
- Scalable production-ready applications.
- Can automate traditionally complex and time-consuming repetitive tasks.

Use cases:

- Study protocol Q&A
- Meeting prep assist

→

- Code autocompletion
- Github copilot PR review

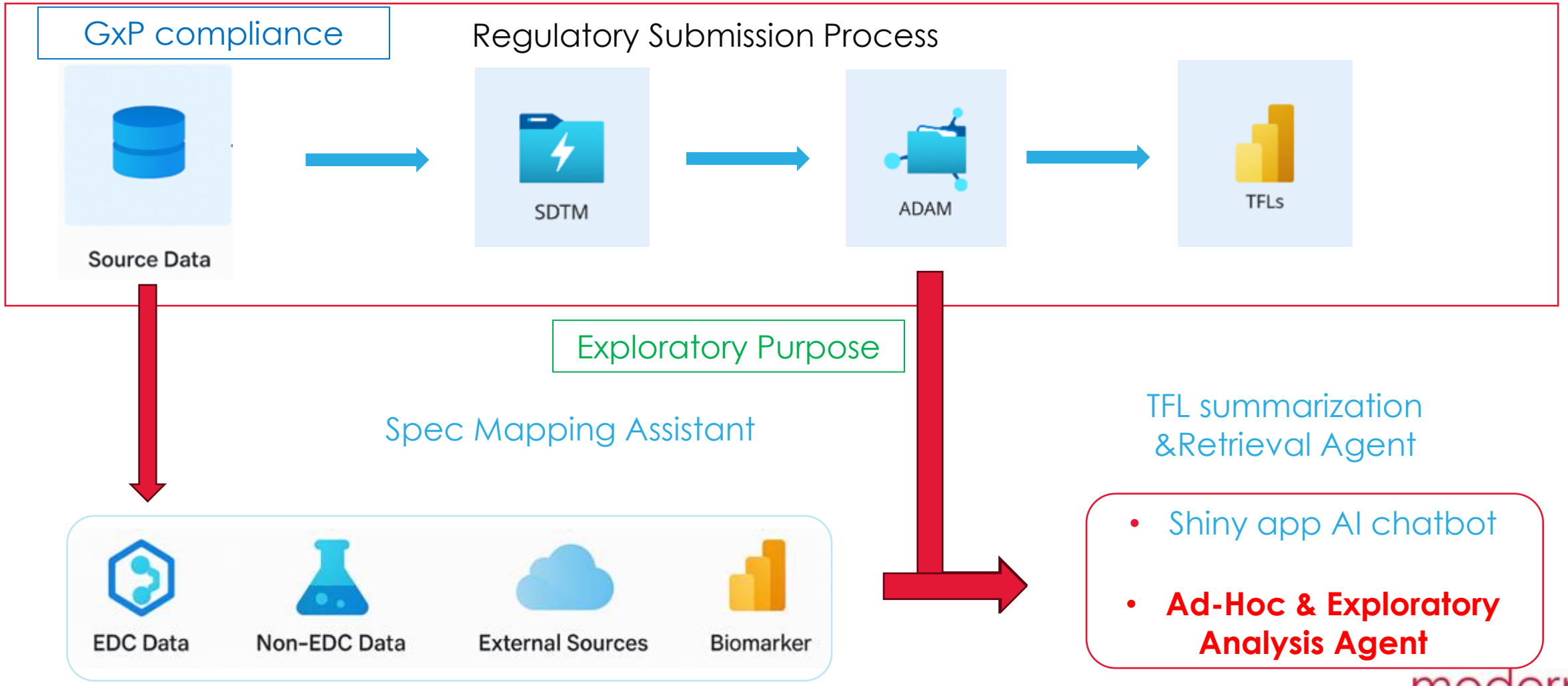
→

- Exploratory Analyses
- TFL generation streamline
- Compliance monitoring

moderna

A Data focused Journey of Exploring Added Business Values in Biometrics

Plan, Retrieve , Execute to Enable end-to-end statistical and analytical automation



Clinical data is abundant, but insights are often delayed

Are clinical benefits associated with baseline characteristics or biomarkers?

What's the time elapse of a safety signal since the most recent dose?

- Allocate resources and identify dataset
- Dataset retrieval
- Statistical analysis
- Visualization
- QC
- Results interpretation

Days to weeks

Result

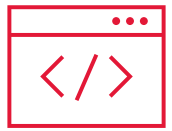
Traditional analysis

AI Agent

AI-powered Analysis

Result

+



Code

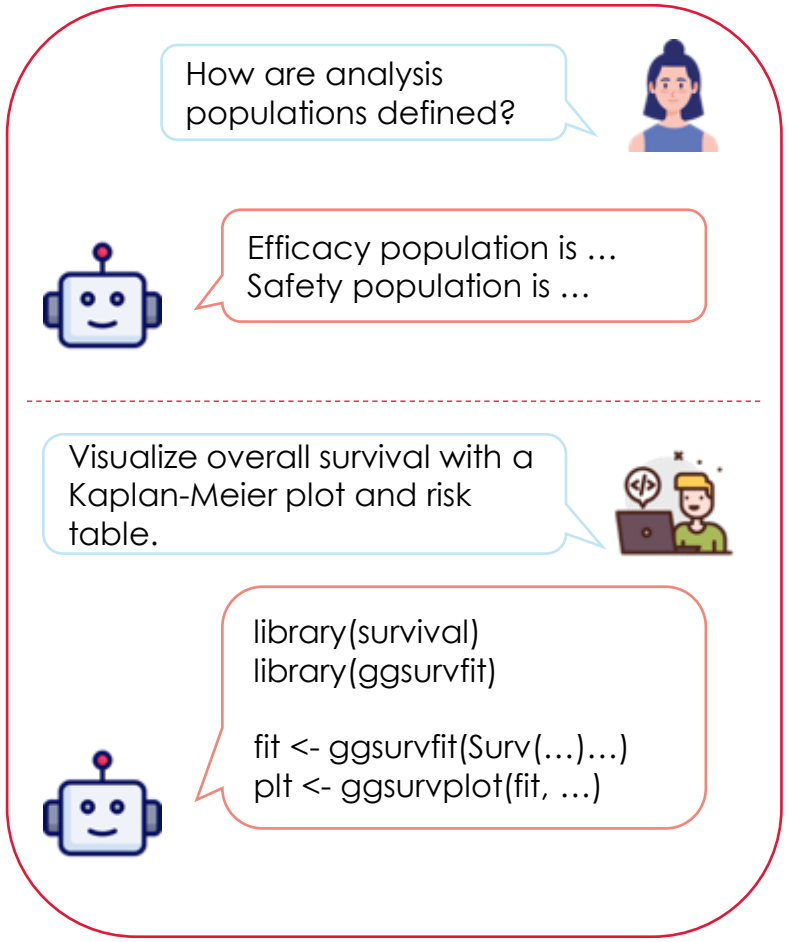
+

Human in the loop

AI has the potential to drastically compress analysis timelines and enable faster, better decision-making.

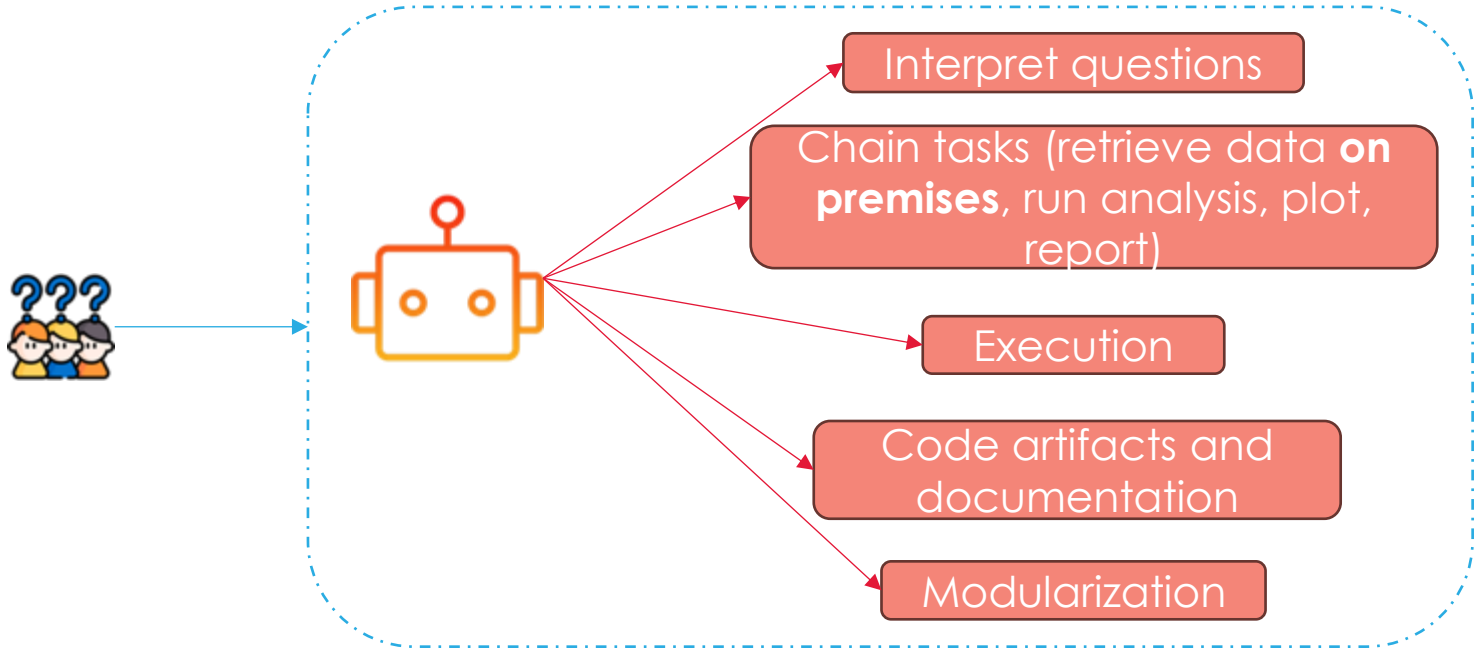
moderna

Smarter, faster clinical workflows with AI agents



AI Agents: autonomous systems that integrate and execute workflows

- **Understands analysis requests in natural language**
- **Retrieves relevant datasets locally and runs scripts automatically**
- **Produces traceable outputs within minutes**
- ***Modules and Adaptability***



Isolated, embedded
+
human instructions and executions

End-to-end pipelines in minutes
+
Human governance

moderna

AI agent to streamline data analysis: architecture overview

Database preparation

- Automated dataset annotation
- Database creation with semantic labels

Multi-LLM orchestration

- Question interpretation & classification
- Database query and records retrieval
- Analysis & visualization
- Result interpretation

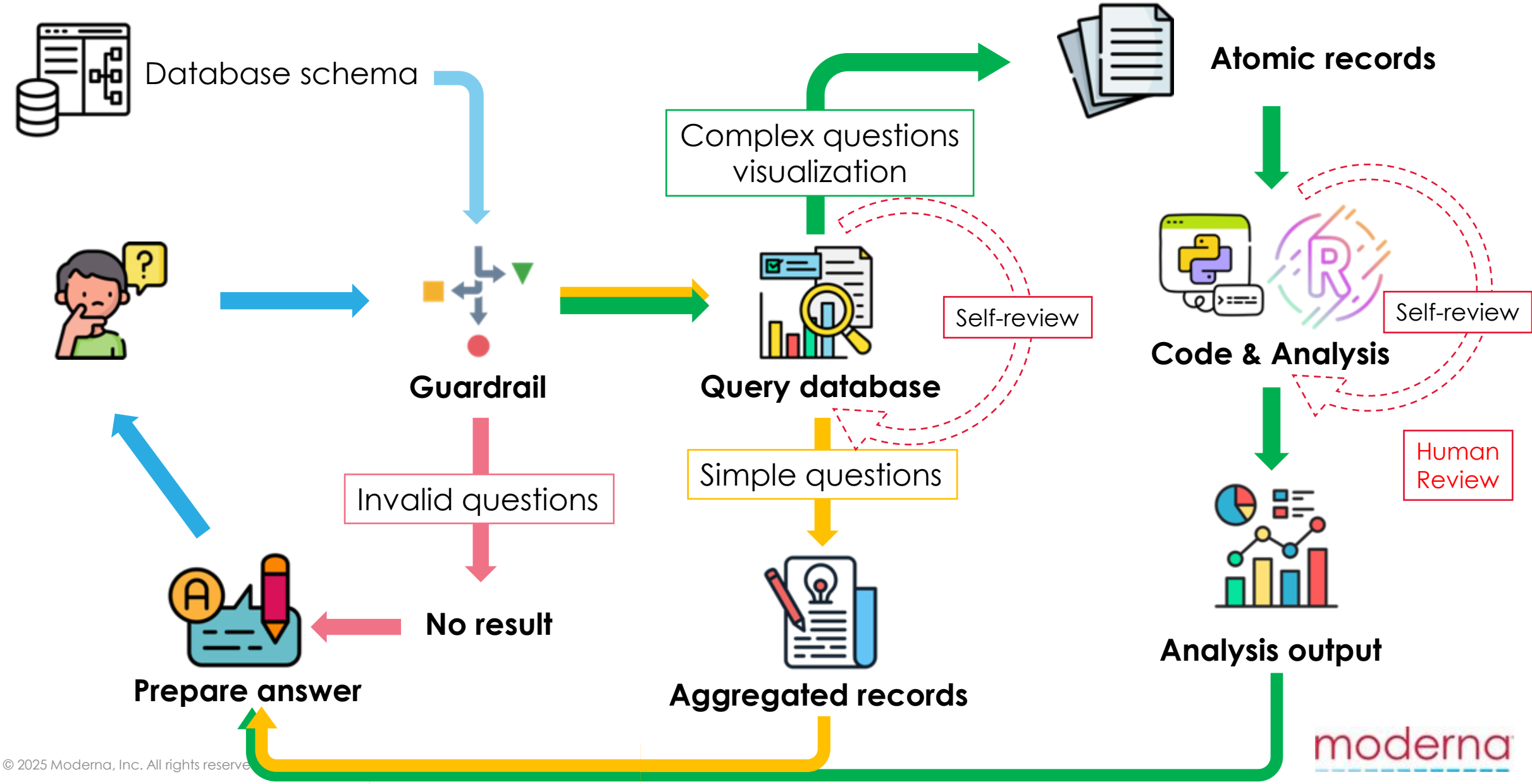
Integration with R/Python

- Alignment with standard workflows
- Trackability & reproducibility

LangGraph + API Key

- Task chain and workflow definition
- Modularity for plug-and-play

AI agent can answer questions with varying complexity



Accelerating clinical insights: from days to minutes

Q: In study X, is progression free survival associated with treatment and age group?

Days:

- ✓ Programmer
- ✓ Data analyst
- ✓ Biostatistician
- ✓ Clinical…

- ? Population flag
- ? Source data
- ? Statistical models
- ? Plot…

Minutes:

Demo

moderna

Demo

Clinical Data Analytical Tool (Internal Preview)

Piloted by ***Statistical Innovations and Data Analytics***. This tool aims to assist clinical data analysts in generating database queries and performing data analyses using natural language prompts. It leverages a combination of LLMs, a graph-based workflow engine, and a clinical research database to facilitate efficient data exploration and analysis.

> How it works

> Tips

Select study:

mRNA-9999-Oncology-DEMO1 mRNA-9999-Oncology-DEMO2

For available fields, please refer to the interactive graph database schema to the right.

Analysis (if applicable) language:

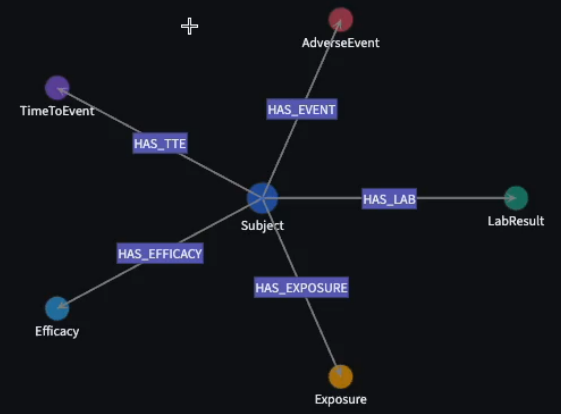
Python
Good for exploratory analysis

R
Good for statistical modeling

Analysis code will be generated in the selected language.

Start New Chat and Clear History

Graph Database Schema for Clinical Data



> Chat History

Ask me something about the study...

> Context information:

Only **synthetic** data is used for this demo.

Future work

Scalability

- Improve automation in database preparation for various source data
- Workflow optimization for robust and reproducible analysis
- Unit test, engineer optimization with Digital team

Compliance

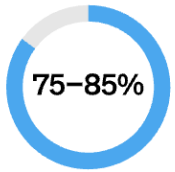
- Data governance, compliance, Validation and Audit Tails
- Standardization of code adherent to company policy
- Human in the loop on every steps

Adaptability

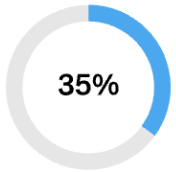
- Continue modularization to expand adaptability
- Integrating latest tools and approaches e.g. MCP , Claude code SDK
- Feedbacks from Users and RL/Fine Tune with local model

Thank you

Work will transform

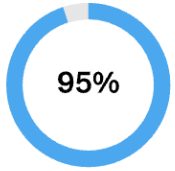


Agentifiable workflows
Processes in life sciences where key activities can be augmented or automated by agents

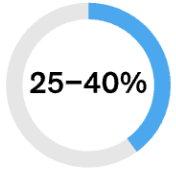


Lower-complexity agents
Agents that require minimal coding, with mostly business-driven build and oversight and rewiring mechanism for adoption

Roles will change



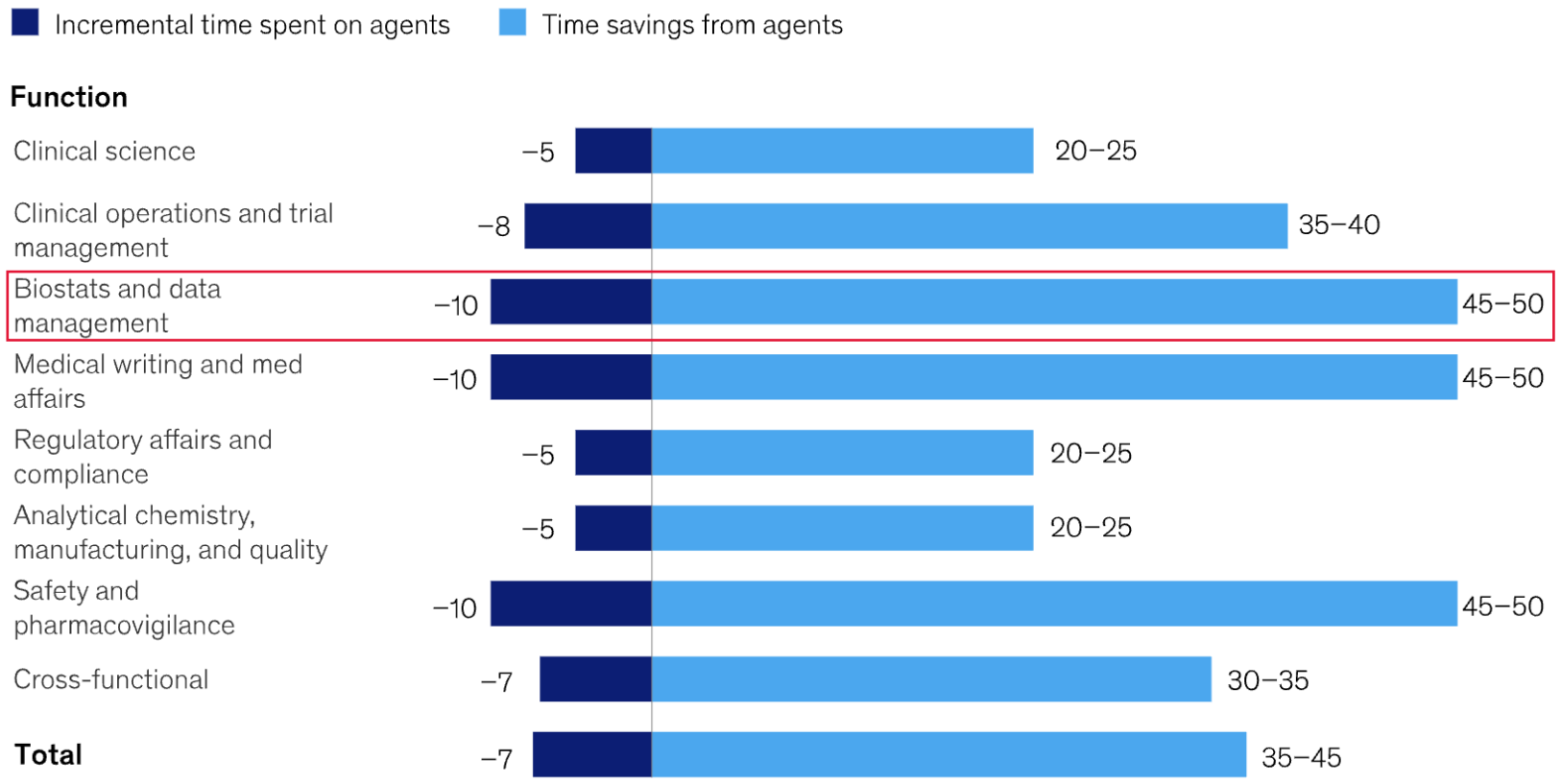
Roles with AI sidekicks
With agents as integrated teammates, employees' impact for patients can be amplified



Freed-up enterprise capacity
Assigning agents to tasks where they excel (and humans don't) can free up time for where humans' unique abilities are needed most

Health data science draw benefits from agent deployment

Work hours shifted by agentic workforce, % of function capacity



McKinsey & Company

CDISC standardization

Statistical programming & analysis

TFL reporting

Data integration & management

…

moderna