What matters most: AI × data driven R&D for corporate innovation

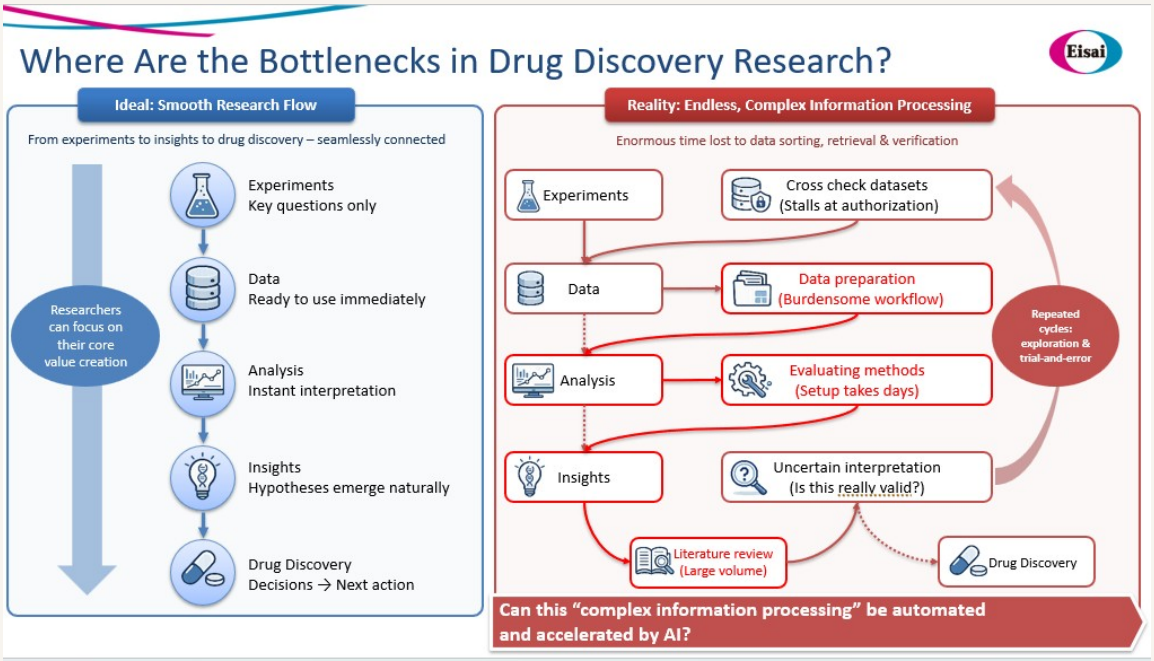
AI-driven drug discovery: data integration and talent development as key drivers

Eisai Co., Ltd.

ELSEVIER

Advancing human progress together

Eisai Co., Ltd. operates under its “hhc (Human Health Care)” corporate philosophy, which places patients and consumers at the center of healthcare and pursues business activities by enhancing their benefits. Within the DHBL Human Biology Creation Hub Integrated Data Science Human Biology Data Ecosystem division - which places strong emphasis on Human Biology - genomic and clinical data are integrated and analyzed using AI and machine learning to support a broad range of activities from drug discovery hypothesis generation to clinical trial design and lifecycle management strategies. Against this backdrop, two major challenges have emerged: keeping pace with the rapid evolution of AI technology and simultaneously developing talent with deep expertise in both AI and drug discovery.



The “bottleneck” in drug discovery and R&D is not simply a lack of data. In most cases, it stems from situation where data “exists but cannot be used,” or from processes that are too complex to complete - spanning data acquisition, AI-driven analysis, and decision-making. Even when the necessary information and data already exist, the sheer volume, the high degree of specialization and fragmentation makes manual processing impractical. As a result, researchers are observed to spend a disproportionate share of their time on “information processing” rather than on scientific inquiry.

Challenges on the research floor: “Hidden burdens”

- Data access and preparation burdens
- Analytical tools and environment setup
- Literature search

Eisai's Initiatives

A defining feature of Eisai's approach is that researchers input what they want to accomplish and the AI is designed not merely to be "built" but to be actively used in day-to-day practice.

(1) Automation of information processing: AI supports from code generation to analysis execution

Drawing on the concept of the AI Scientist / AI Agent, when a researcher inputs what they would like to accomplish, the AI autonomously handles data retrieval, analysis code generation, and visualization - realizing "automation and acceleration." Furthermore, by incorporating a Human-in-the-Loop framework - in which humans verify, correct and approve outputs - the system is designed to operate on the assumption that hallucination (the generation of plausible but factually incorrect outputs) may occur. In practice, Eisai has adopted "Biomni" a customizable, general-purpose biomedical AI agent developed by Stanford University - to autonomously automate information processing. The key point is that AI does not merely execute analyses on behalf of researchers - it also maintains a structure in which researchers can independently evaluate the validity of outputs.

- Rapid adoption of emerging technologies
- AI that is merely "built" will not be adopted in practice - ensuring researchers "actually use it" is the priority
- Deploy even rough versions in the field and refine practical utility: company-wide rollout → feedback → iteration
- Building a co-creation environment for drug discovery researchers and AI developers

(2) Ensuring reliability: Enhancing "reproducibility" grounded in know-how (domain expertise) and governance

General-purpose AI alone is insufficient, it is necessary to codify company-specific organizational, implicit knowledge and embed it into AI systems. At Eisai, domain expertise is added as Know-How files, enabling AI to apply specialized knowledge appropriate to each context. Hallucination - the generation of plausible but factually incorrect outputs — is identified as the greatest risk in AI utilization and the system is designed with this risk acknowledged as inevitable, controlling it through prompt engineering and structural mechanisms.

By embedding researchers' knowledge into AI in a "reusable form," the system achieves high-accuracy, reproducible analysis. In addition, for the scientific literature indispensable to research, LLMs (Large Language Models) are applied to the time-intensive tasks of reading, comparing and summarizing large volumes of publications — enabling high-relevance literature retrieval, query refinement and search path optimization, and connecting evidence verification to interpretation and judgment.

Conclusion

Key Takeaway:

Substantial automation of analytical work
Reducing the operational burden on researchers, enabling focus on interpretation and judgment, freeing up time for iterative experimentation, and improving overall efficiency

Improved evaluation, reproducibility, and reliability
Human verification (Human-in-the-Loop) + domain expertise (Know-How) + intentional design (Design) → evidence-based outputs

Connection to decision-making
Moving toward seamless operations from data acquisition → analysis → visualization → the next action

People in this project:

Keishi Akada
DHBL Human Biology Creation Hub
Integrated Data Science,
Human Biology Data Ecosystem
Principal Scientist

"It is through the co-creation of drug discovery researchers and AI developers that a new culture of value creation — AI-driven drug discovery — is born."