

Challenges of Biomedical Data Science for Women's Health

Jason H. Moore, PhD, FACMI, FIAHSI, FASA
Chair, Department of Computational Biomedicine
Director, Center for Artificial Intelligence Research and Education
Cedars-Sinai Medical Center
Los Angeles, CA

jason.moore@csmc.edu

Learning Health Systems

A system in which science, informatics, incentives, and culture are aligned for continuous improvement of the healthcare delivery experience (IOM 2015).

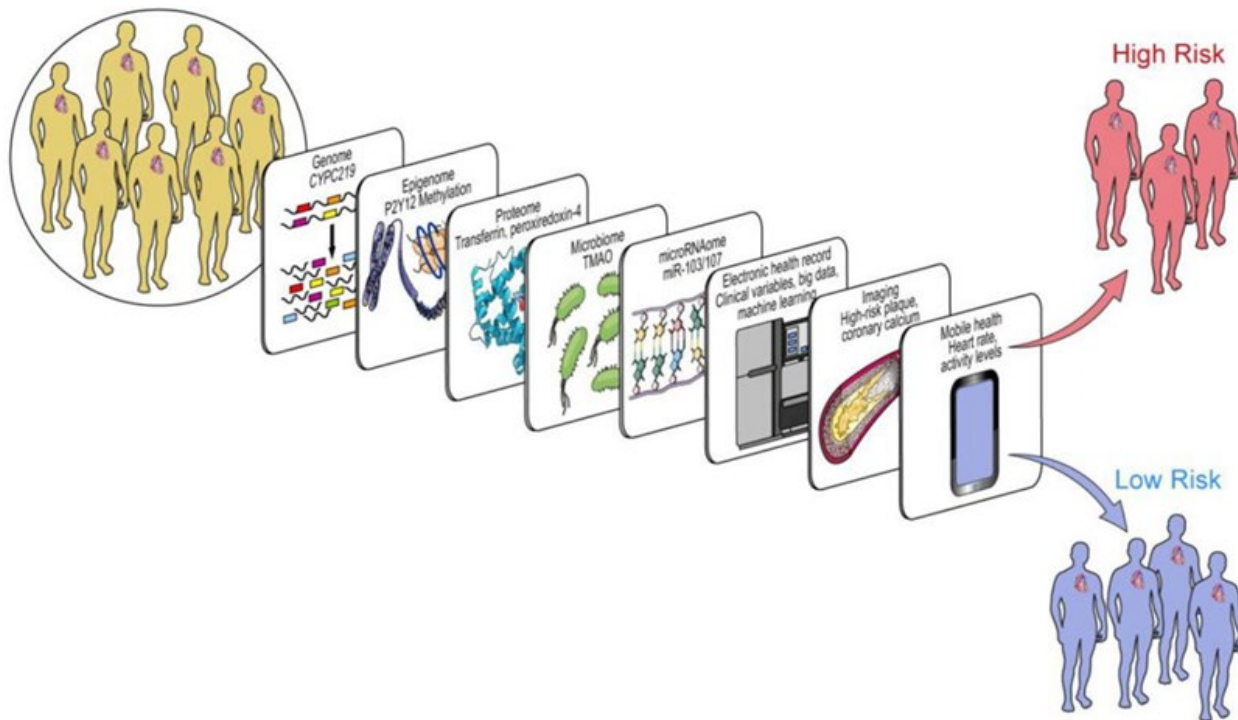


Challenges of Learning from Data

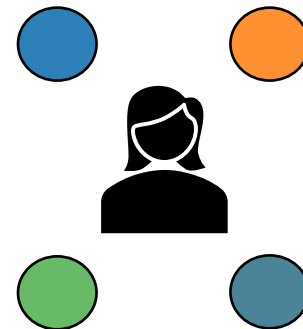
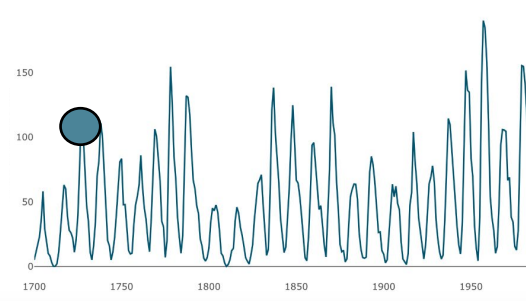
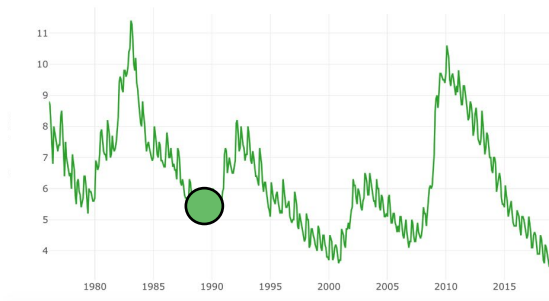
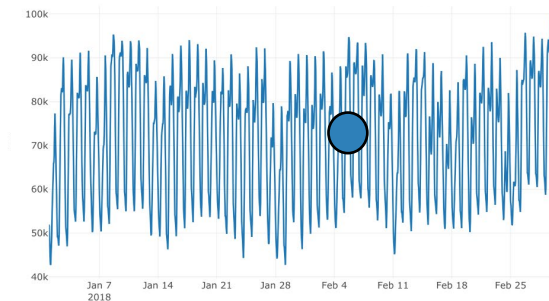
Data are....

1. Expensive
2. Often static
3. Messy
4. Difficult to integrate
5. Private
6. Difficult to analyze
7. Difficult to operationalize

Data are expensive



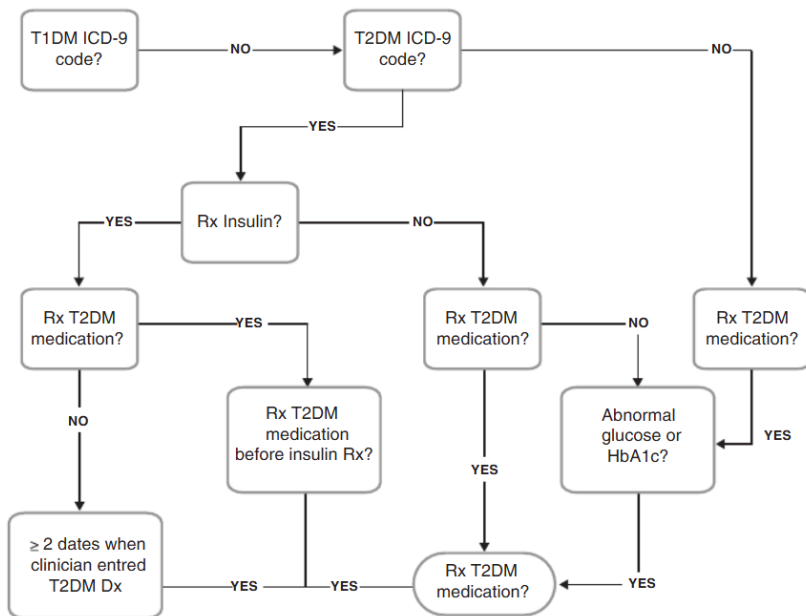
Data are often static



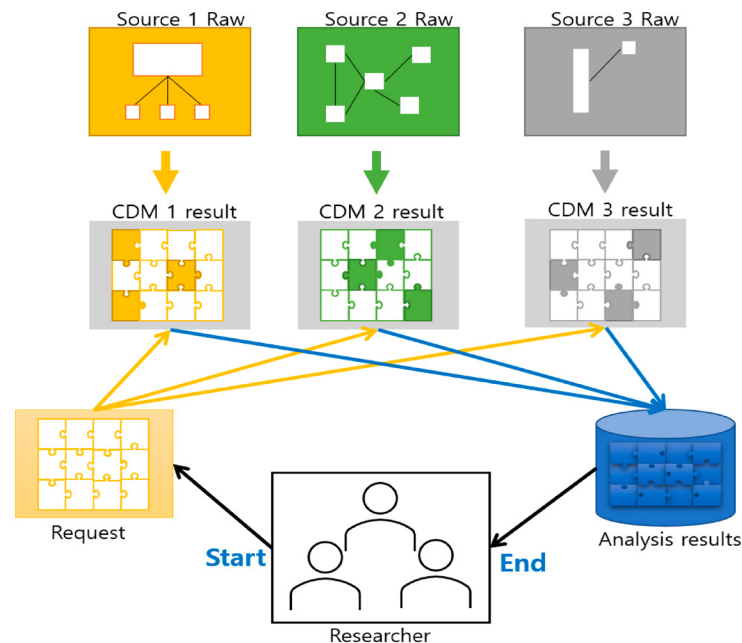
[illegible]

Data are difficult to integrate

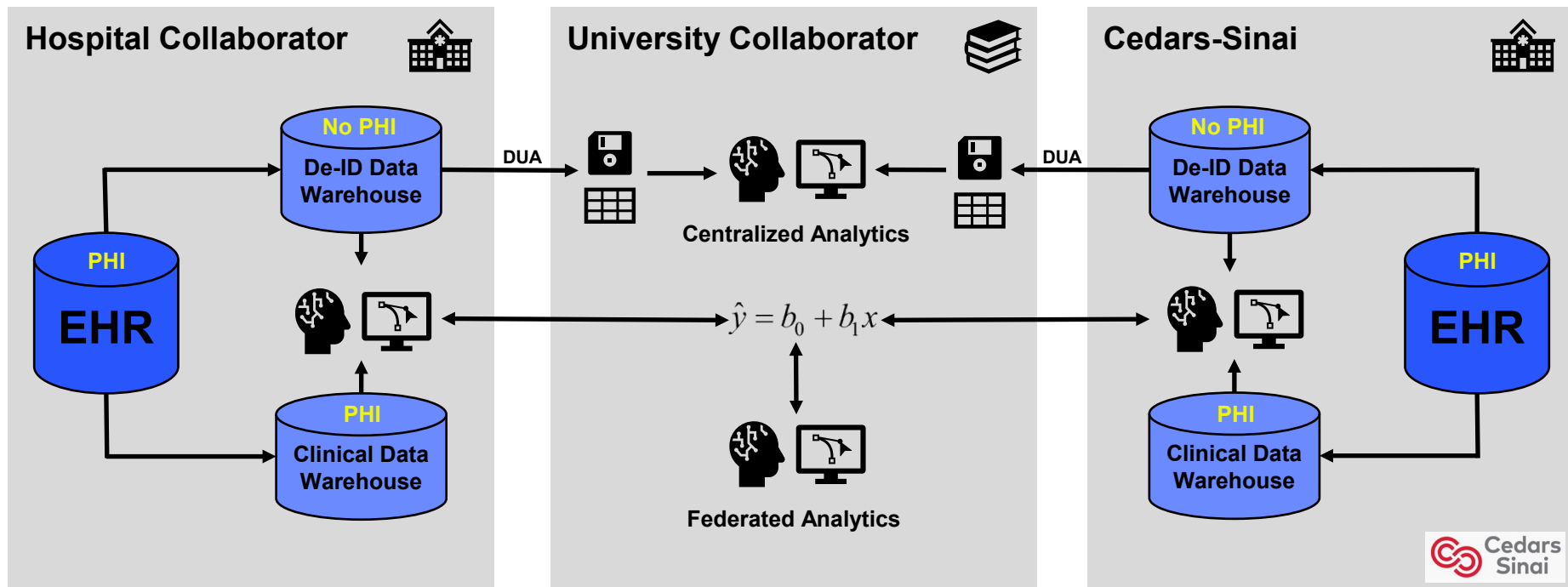
Phenotyping algorithms



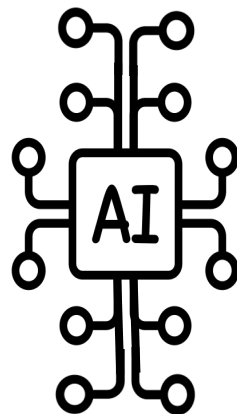
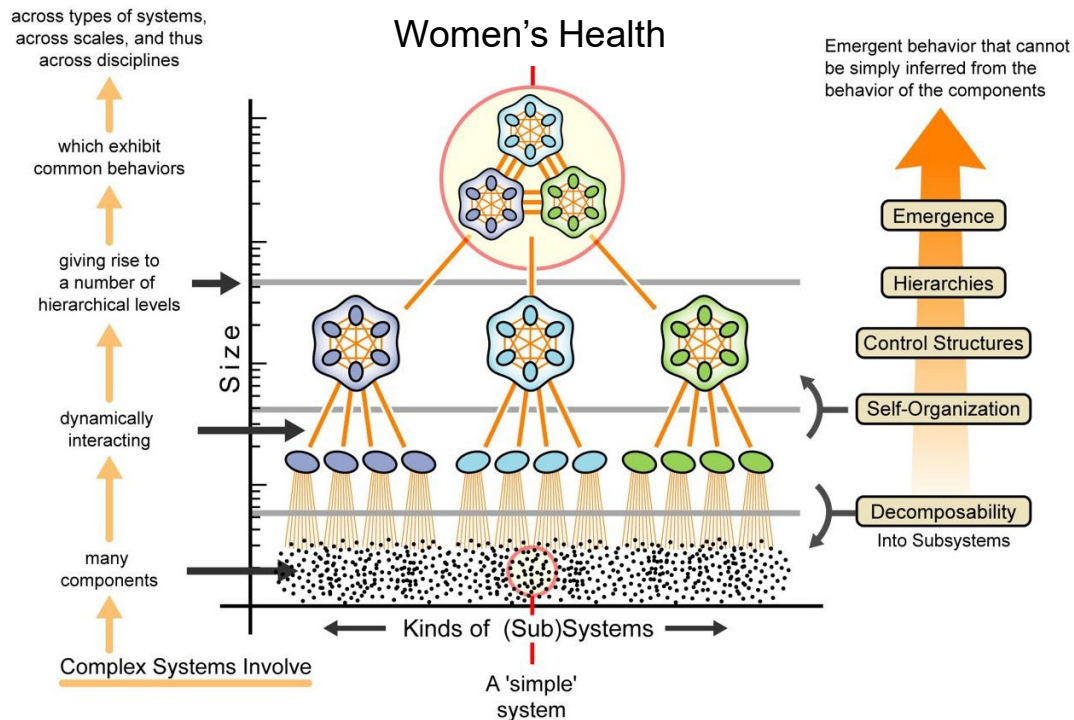
Common data models



Data are private – Centralized vs. Federated Analytics

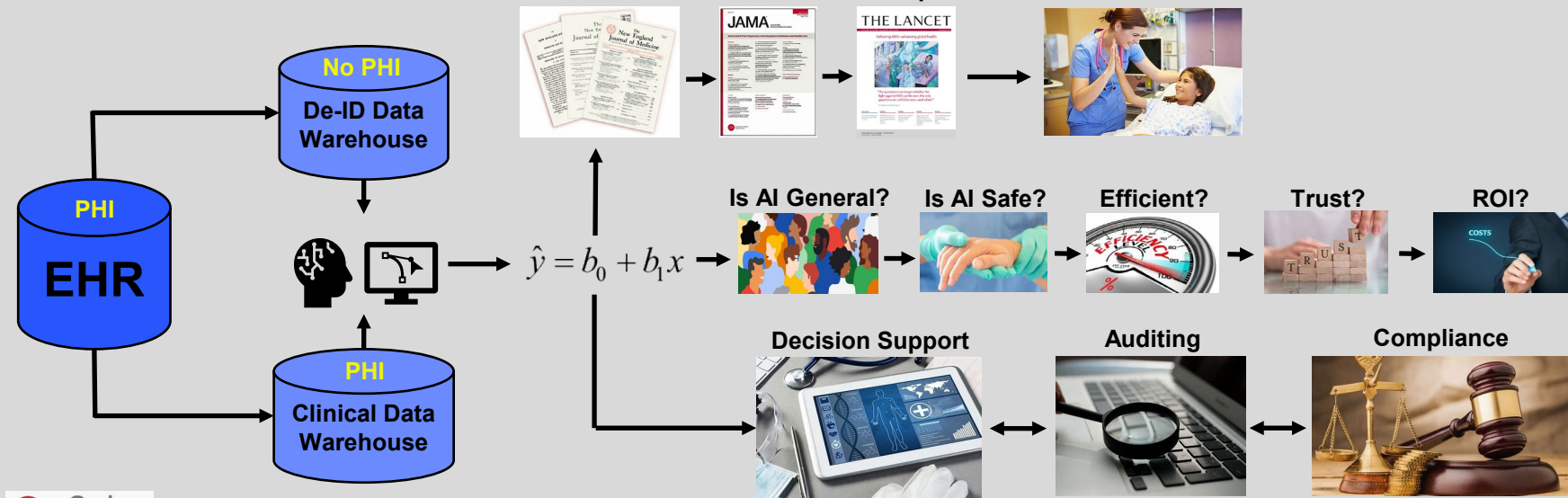


Data are difficult to analyze



Data are difficult to operationalize

Cedars-Sinai



**Easy Data
Access**

**Easy Data
Cleaning**

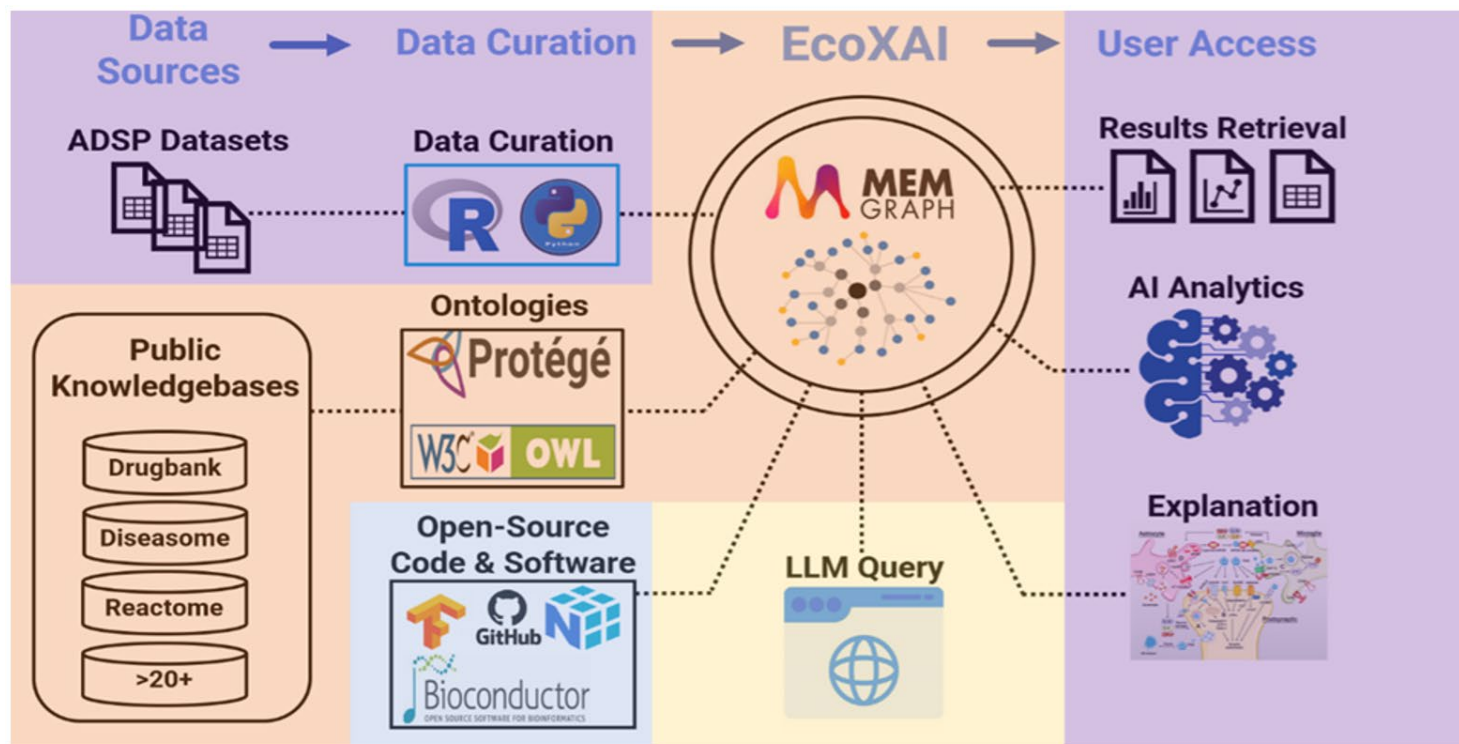
**Easy Data
Analytics**

**Easy Model
Evaluation**

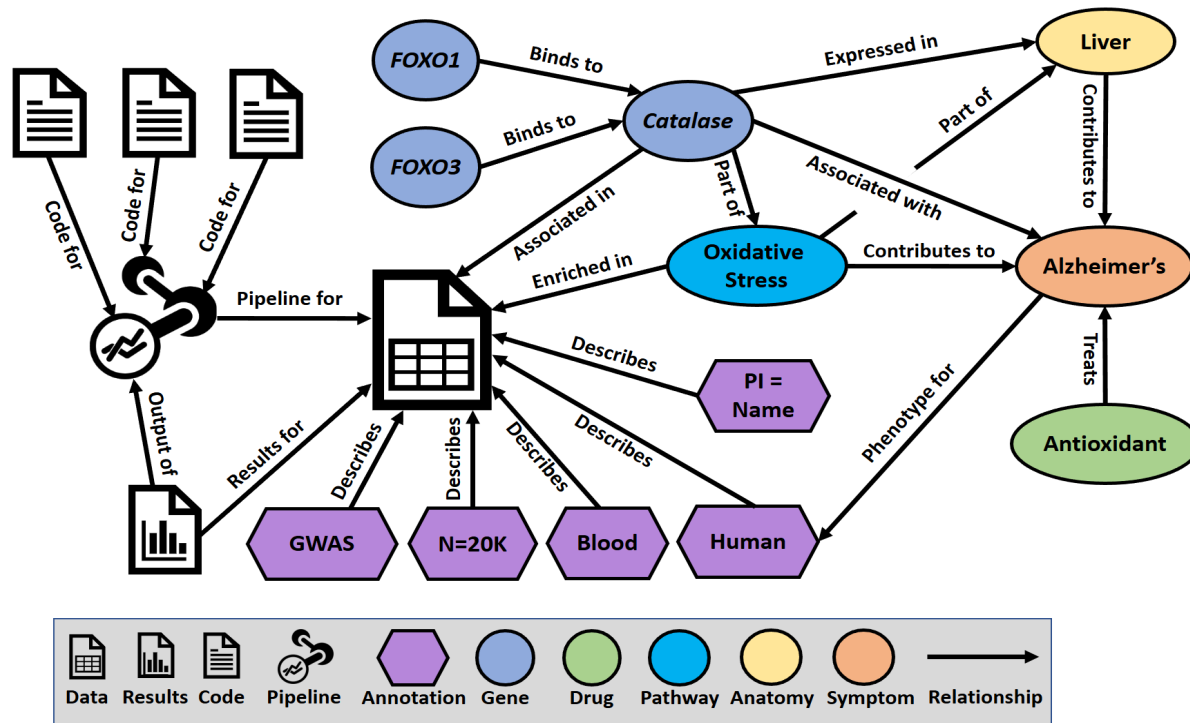
**Easy Model
Deployment**



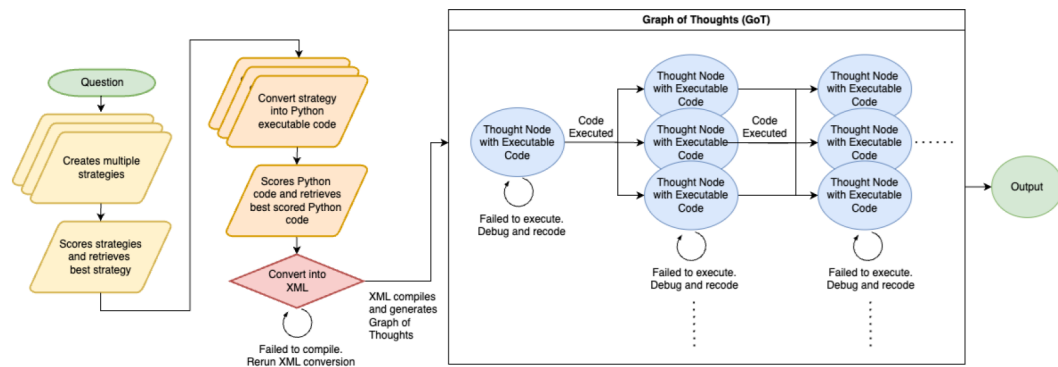
An Ecosystem for Explainable Artificial Intelligence (EcoXAI)



An Ecosystem for Explainable Artificial Intelligence (EcoXAI)



Accurate use of AI to query EcoXAI using natural language



Bioinformatics, 2025, 41(2), btaf031
<https://doi.org/10.1093/bioinformatics/btaf031>
 Advance Access Publication Date: 22 January 2025
 Applications Note

Bioinformatics
 OXFORD

Databases and ontologies

ESCARGOT: an AI agent leveraging large language models, dynamic graph of thoughts, and biomedical knowledge graphs for enhanced reasoning

Nicholas Matsumoto¹, Hyunjun Choi¹, Jay Moran¹, Miguel E. Hernandez¹, Mythreye Venkatesan¹, Xi Li¹, Jui-Hsuan Chang¹, Paul Wang¹, Jason H. Moore^{1,*}

¹Department of Computational Biomedicine, Center for Artificial Intelligence Research and Education, Cedars Sinai Medical Center, West Hollywood, CA 90069, United States

Dataset	GPT-4o Mini (%)	RAG (%)	ESCARGOT (%)
Openended 1-hop (508 questions)	4.2	55.5	88.4
Openended 2-hop (450 questions)	4.9	23.1	85.8
True/false 1-hop (560 questions)	60.5	85.2	90.9
True/false 2-hop (540 questions)	59.4	75.6	74.1
Multiple choice 1-hop (498 questions)	58.2	88.8	93.4
Multiple choice 2-hop (419 questions)	61.3	86.4	89.5

**Easy Data
Access**

**Easy Data
Cleaning**

**Easy Data
Analytics**

**Easy Model
Evaluation**

**Easy Model
Deployment**



Acknowledgements

- EcoXAI: Jun Choi, Miguel Hernandez, Nick Matsumoto, Jay Moran, Paul Wang
- R01 LM010098, U01 AG066833
- jason.moore@csmc.edu

