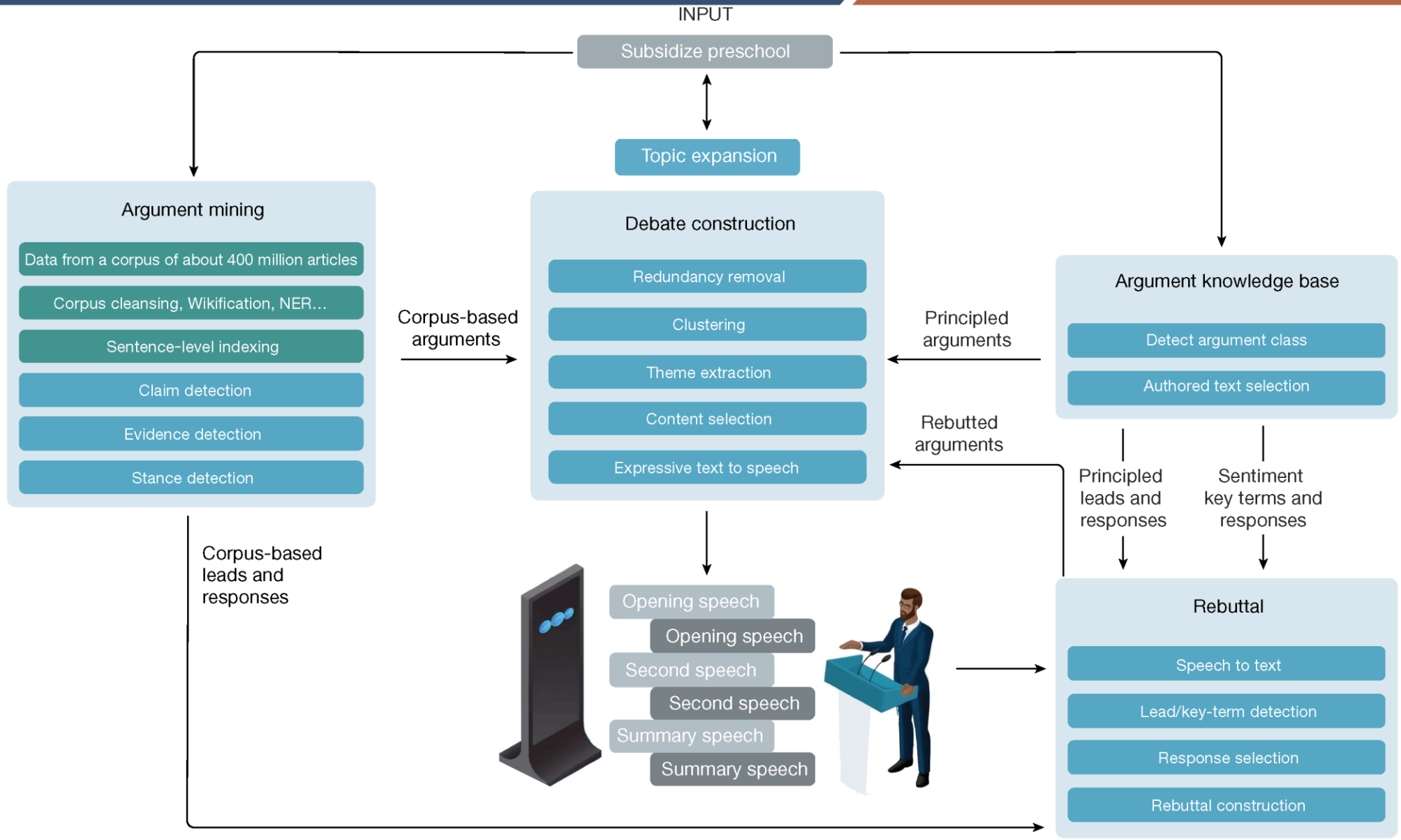
AI-Agents for Bioinformatics and Systems Medicine

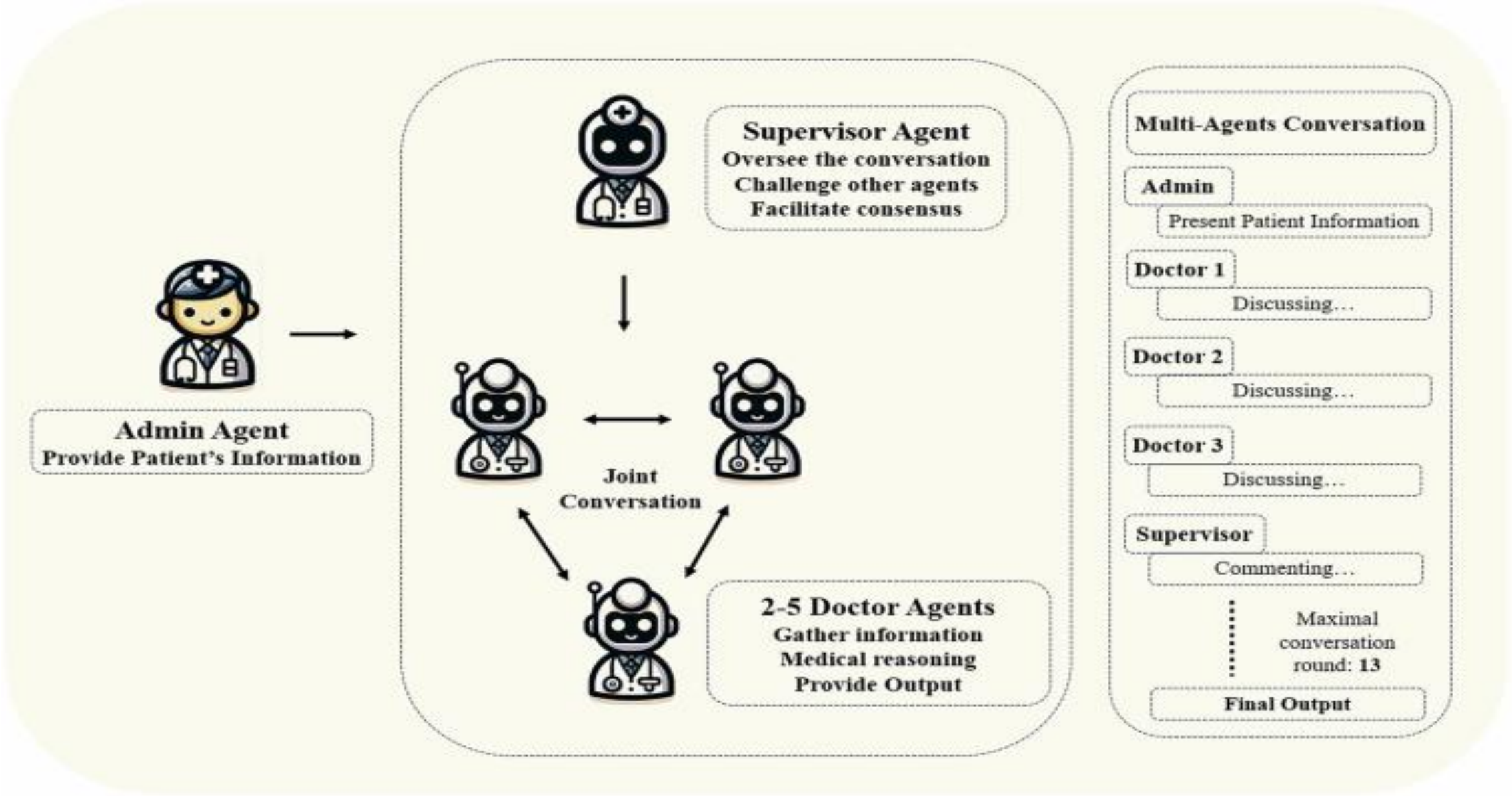
McWilliams School of Biomedical Informatics

Xiaobo Zhou Ph.D.
Fellow of AIMBE and ACMI
Carl V Vartian Professor and Director,
McWilliams School of Biomedical Informatics,
University of Texas Health Science, Houston, TX

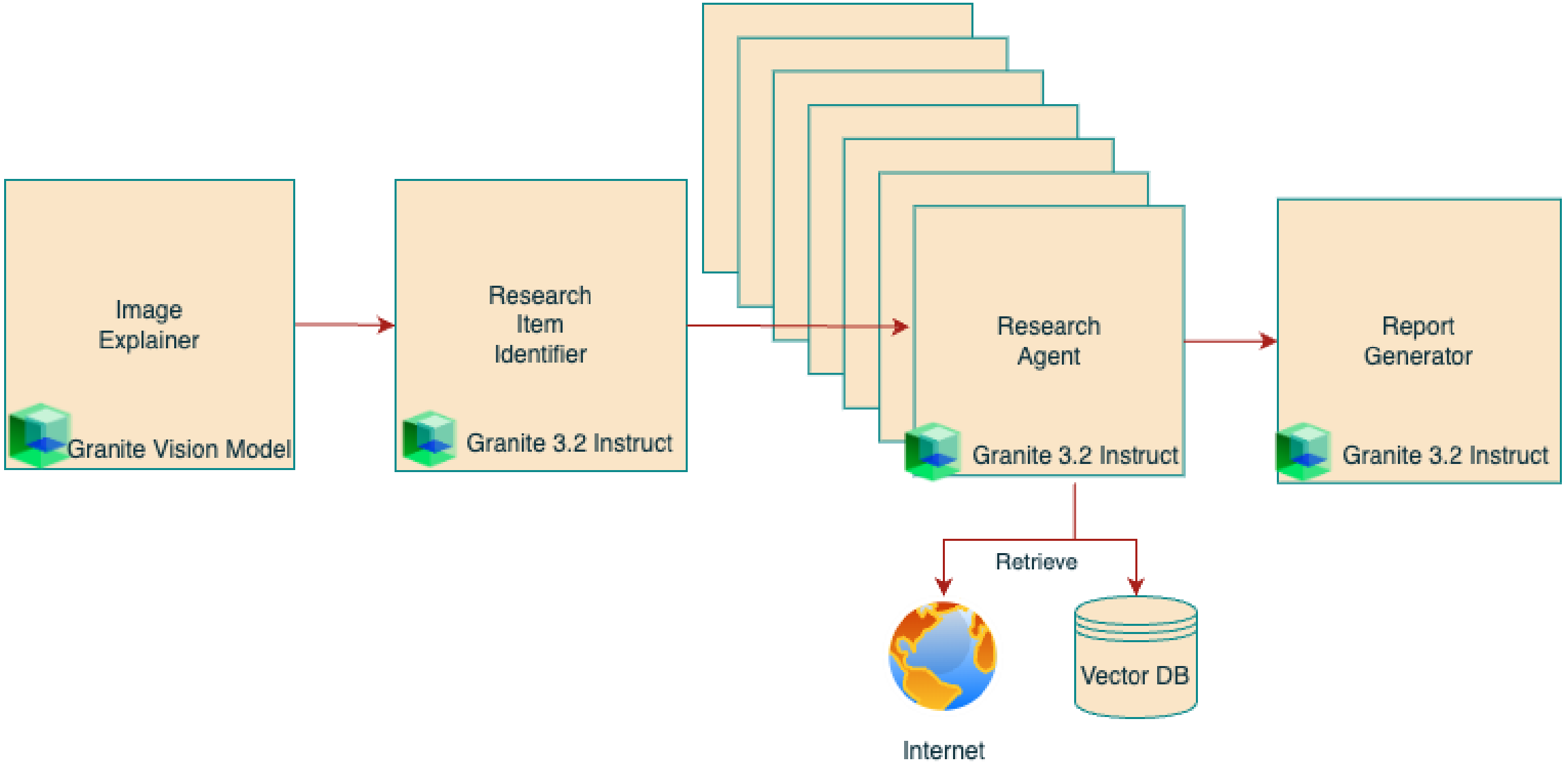
Nature 2021: Autonomous Debating System



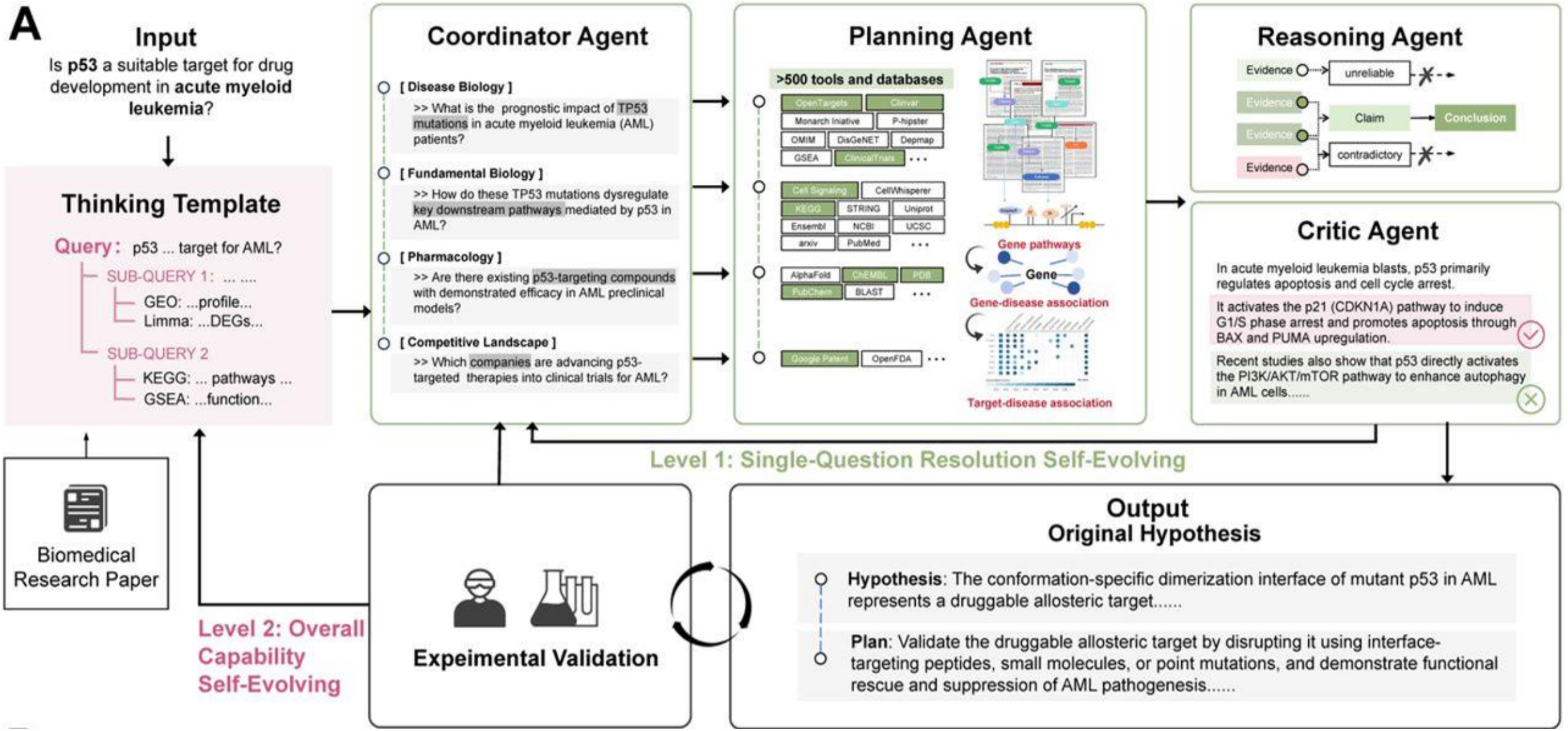
Enhancing diagnostic capability with multi-agents conversational (MAC) large language models



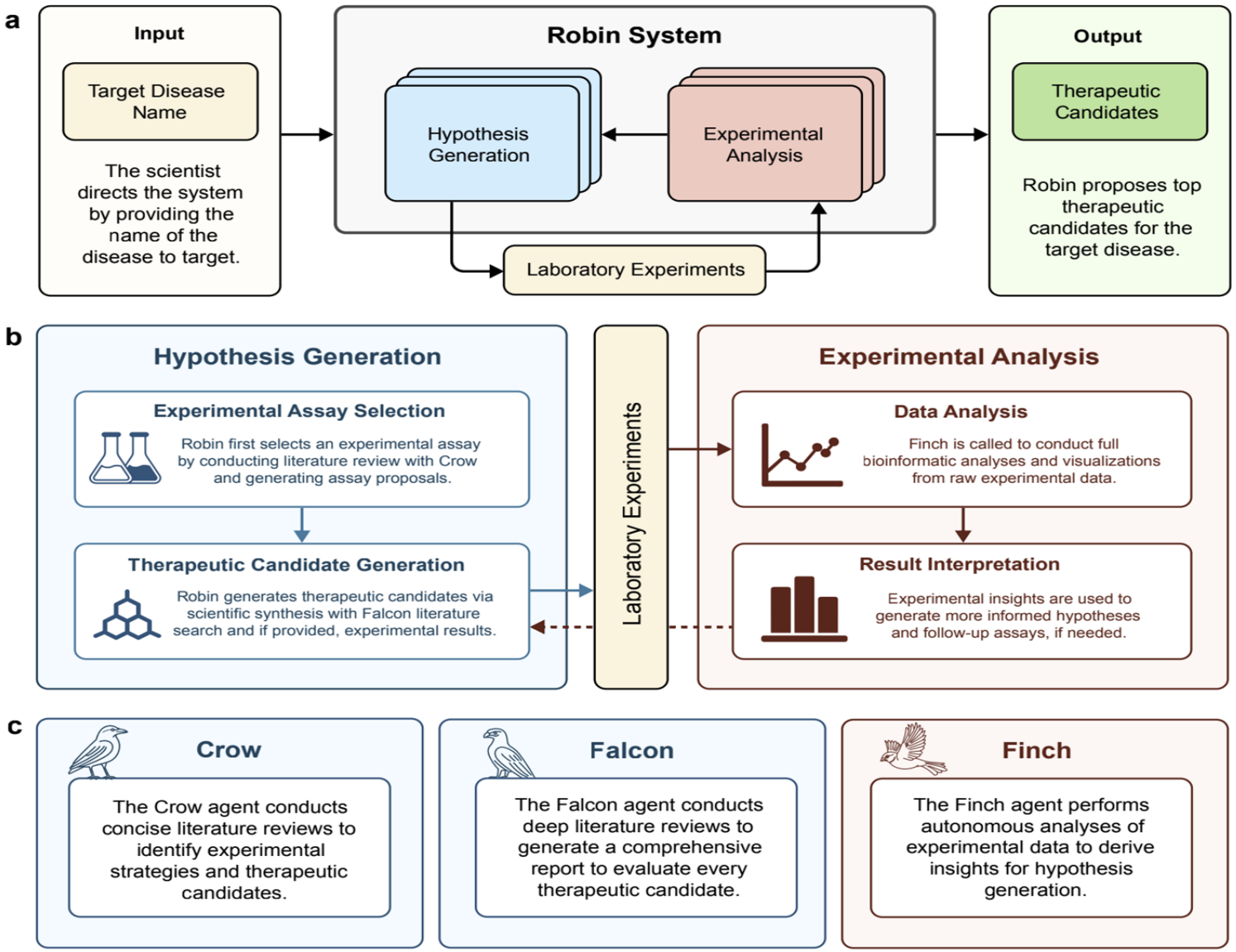
IBM Developer: Build an AI research agent for image analysis with Granite



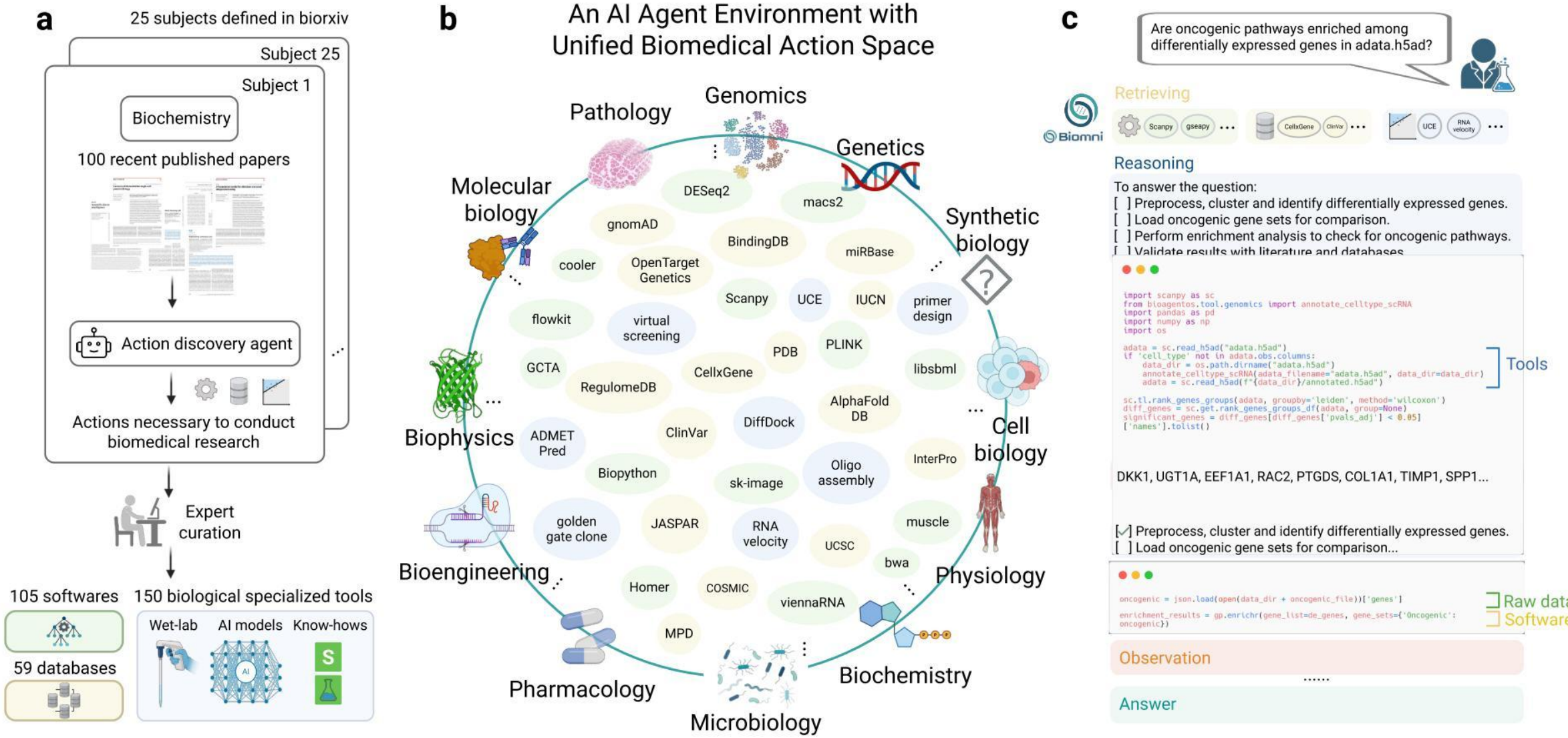
OriGene → self-evolving multi-agent system designed as a virtual disease biologist



Robin → A Multi-Agent AI system for biological discovery

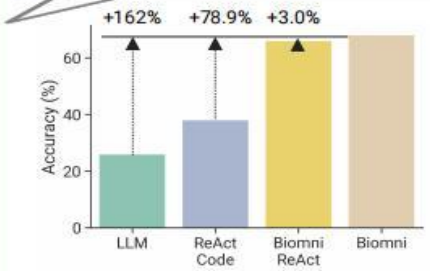


Biomni: A General-Purpose Biomedical AI Agent



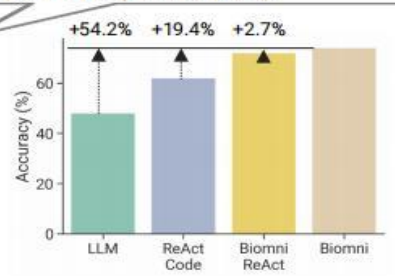
Variant prioritization

I have a list of potential variants for disease X, prioritize the top variant for me.



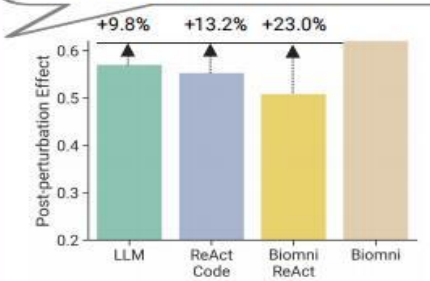
GWAS causal gene detection

I have a GWAS hit for trait XX, there are genes XXX, XXX in the locus, which is the most likely causal gene?



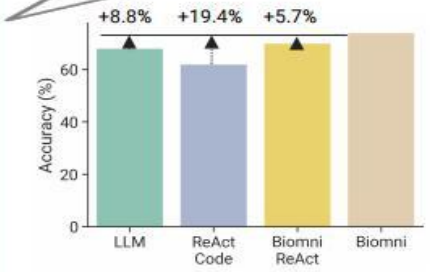
Perturbation screen design

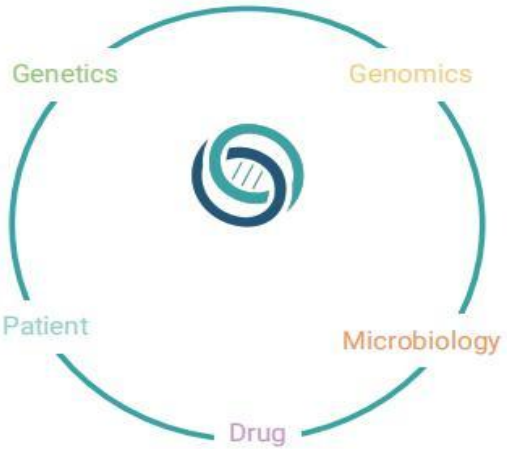
Design a gene panel of 50 genes for my CRISPR screen on T-cell exhaustion.



Patient gene prioritization

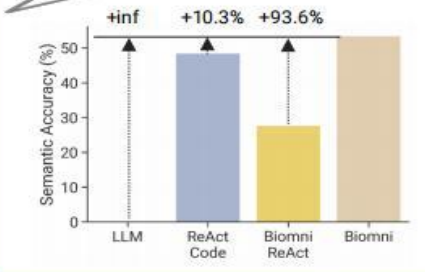
I have a patient with phenotypes XXX, XXX, and WGS returns gene mutations XXX, what is the causal gene?





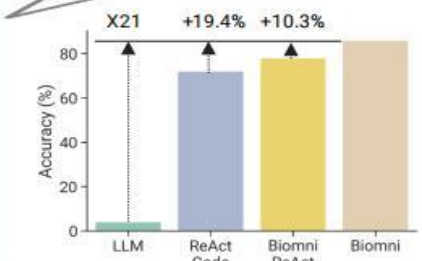
scRNA-seq cell annotation

I have scRNA-seq data stored at XXXX, annotate the cell type for me.



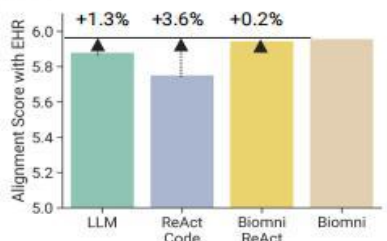
Rare disease diagnosis

I have a patient with phenotypes XXX, XXX, and WGS returns gene mutations XXX, what is the diagnosis?



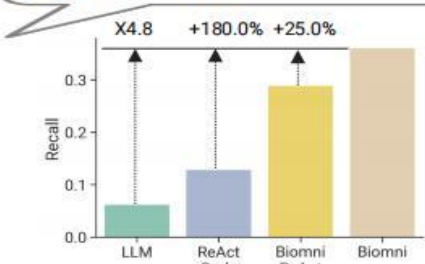
Drug repurposing

I have a rare disease XX, among a list of potential drugs, find the most likely one for repurposing.



Microbiome disease-taxa bioinformatics analysis

I have microbiome raw data stored at XXXX, analyze which taxa is associated with the disease XXX.

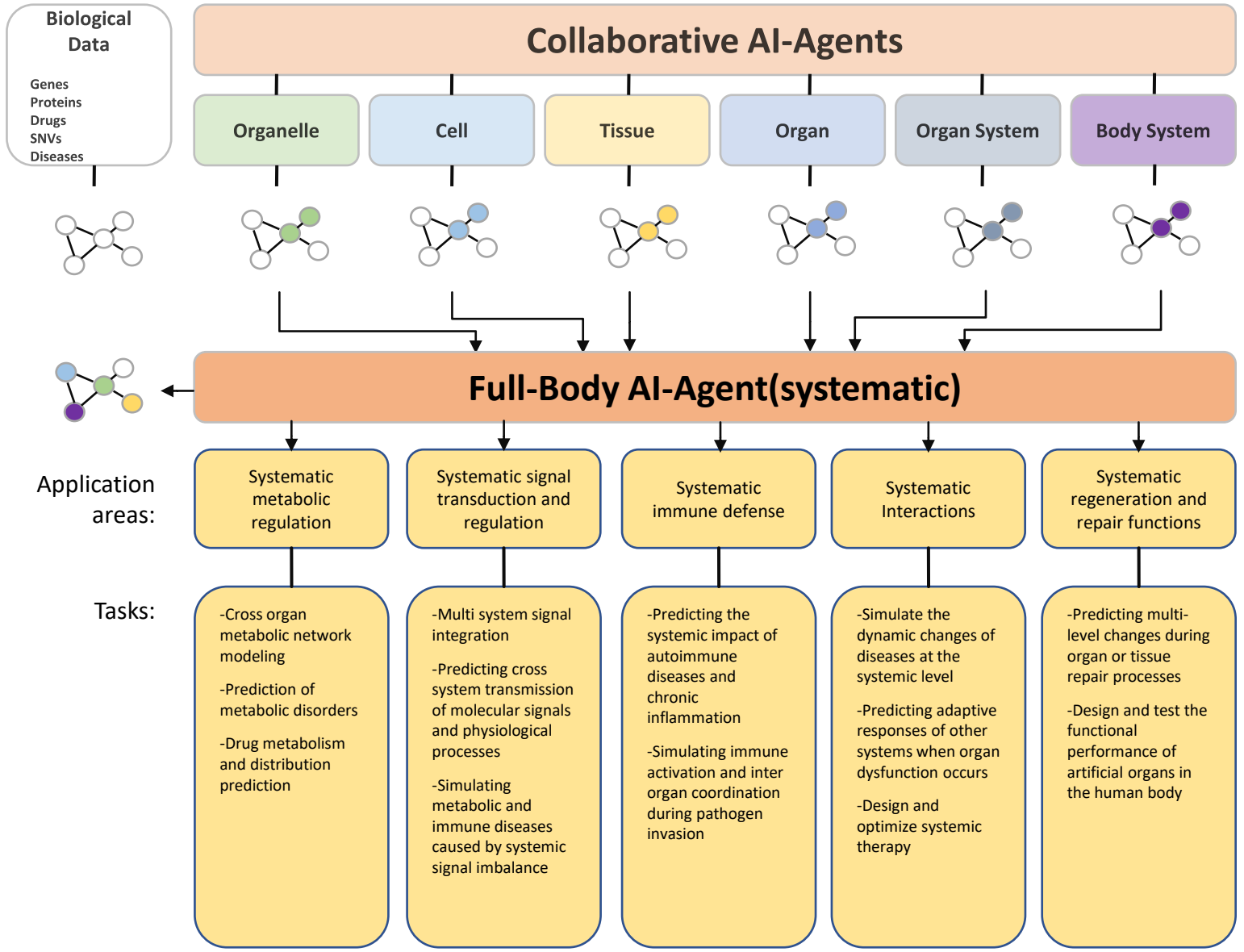


Overview of the Multi-Agent AI systems

	OriGene	Robin	Biomni
General Purpose	Discover novel therapeutic targets using self-evolving reasoning templates	Automate closed-loop scientific discovery (lit. review, hypothesis generation, experiment design and analysis). Simulates entire lab team through multi-agent setup	General purpose biomedical agent that plans and executes biomedical tasks across 25 subfields, uses a code-native reasoning engine that writes and executes code for task-specific pipelines
Workflow Type	Structured reasoning based on curated template	Dynamic multi-agent reasoning loop: literature (Crow), deep synthesis (Falson), data analysis (Finch)	Fully flexible, code-driven workflows, adapts pipelines on-the-fly for diverse task types
Tool access	500+ biomedical databases, ontologies, and pathway tools (ex: GEPIA, OpenTargets, PubMed)	Embedded LLMs that query literature and prompt engineer Finch. Literature agents built off of PaperQA2 which accesses/retrieves literature from local file repositories	150 specialized and curated tools, 105 software packages, 59 biomedical databases

The New Era: Full-Body AI-Agent

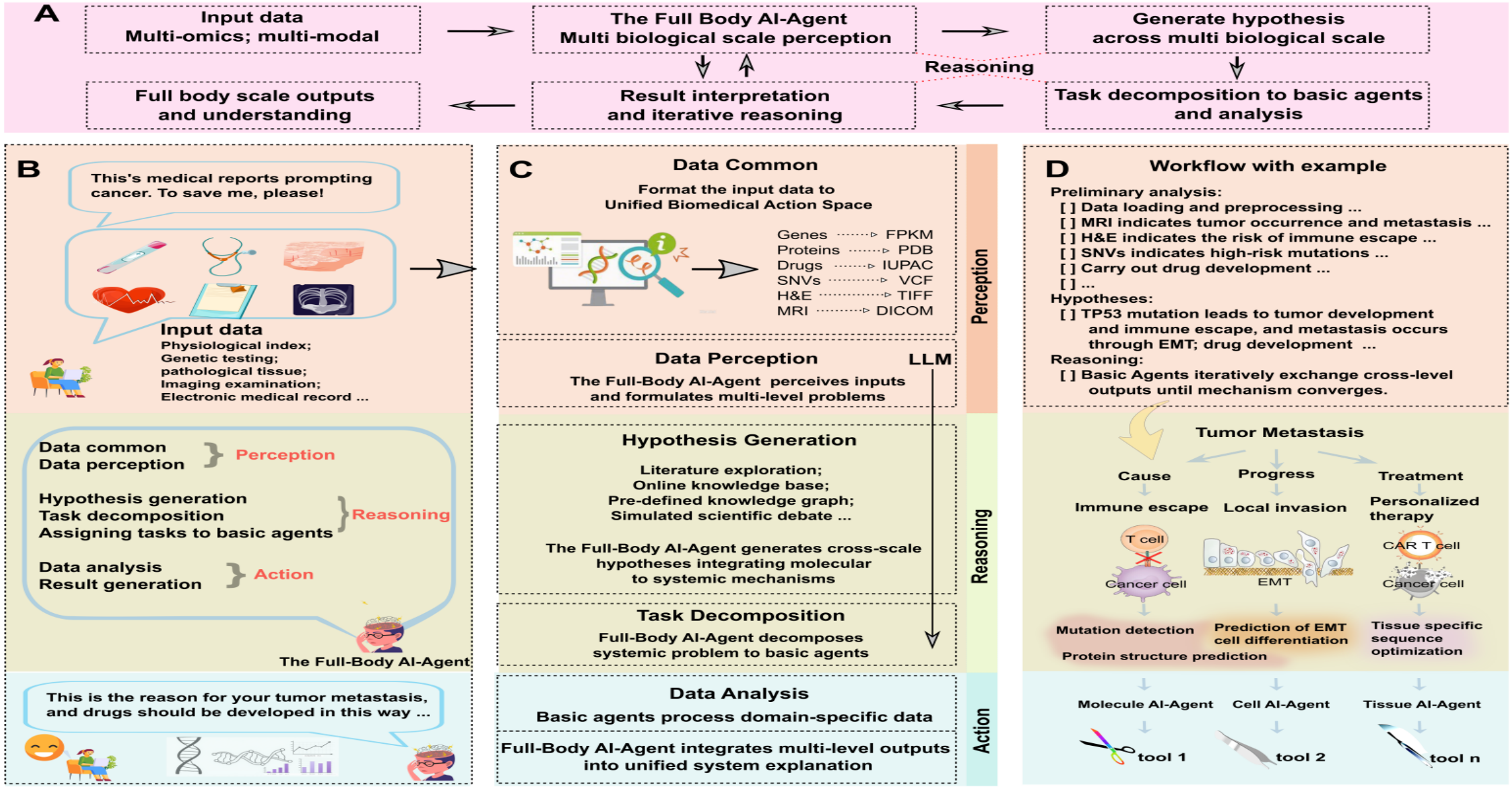
Aoqi Wang



To understand the complex physiological and pathological processes and to address the complexity of human biology through a unified, multi-level framework that bridges molecular, cellular, tissue, organ, and systemic dynamics.

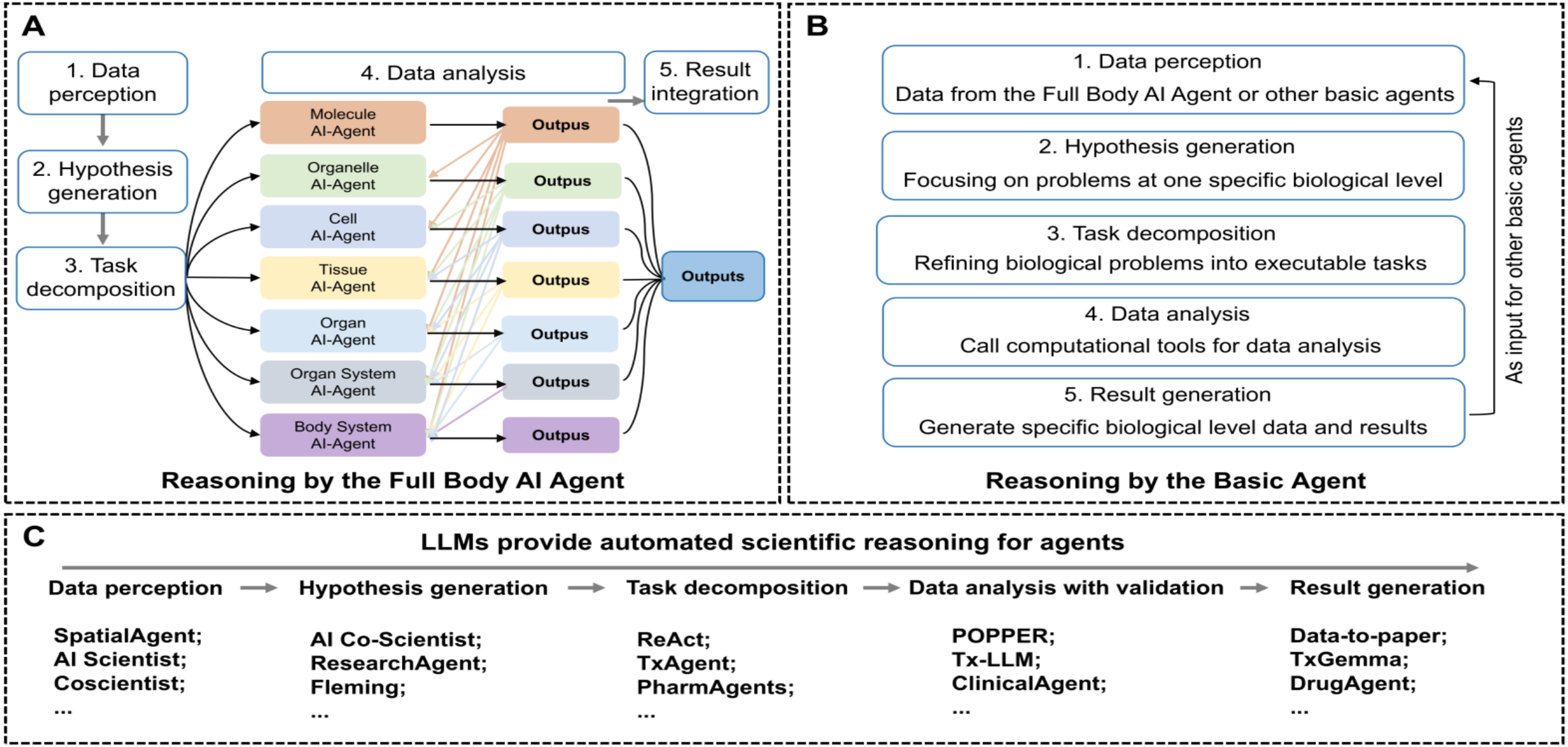
The system emphasizes cross-level collaboration, enabling the study of how molecular changes cascade into cellular behaviors, tissue responses, organ functions, and systemic outcomes.

Schematic of the Full-Body AI Agent for systemic human biology analysis and therapeutic Insight generation



Hierarchical reasoning framework of Full-Body and basic AI Agents

Aoqi Wang



Data

DNA sequence

Protein sequence

Small molecule chemical sequence

Protein structure data

Protein physicochemical data

Protein chemical modification data

Molecule AI-Agent

Application functional areas

Task refinement and execution

Protein structure prediction

-Prediction of protein folding patterns and three-dimensional structures from sequence data

-Simulate changes in protein conformation

Molecular docking

-Prediction of protein-protein interactions

-Prediction of protein complexes and structures

-Prediction of protein and small molecule binding

-Prediction of and nucleic acid binding

-Prediction of the docking molecules based on protein conformation

Post translation modification analysis

-Prediction of post-translational modification sites

-Simulate and predict the effects of PTM on protein structure, function, and stability

-Design precision therapeutic drugs targeting the PTM pathway

Protein functional annotation

-Annotations on the functional domains of proteins

-Classification of protein families

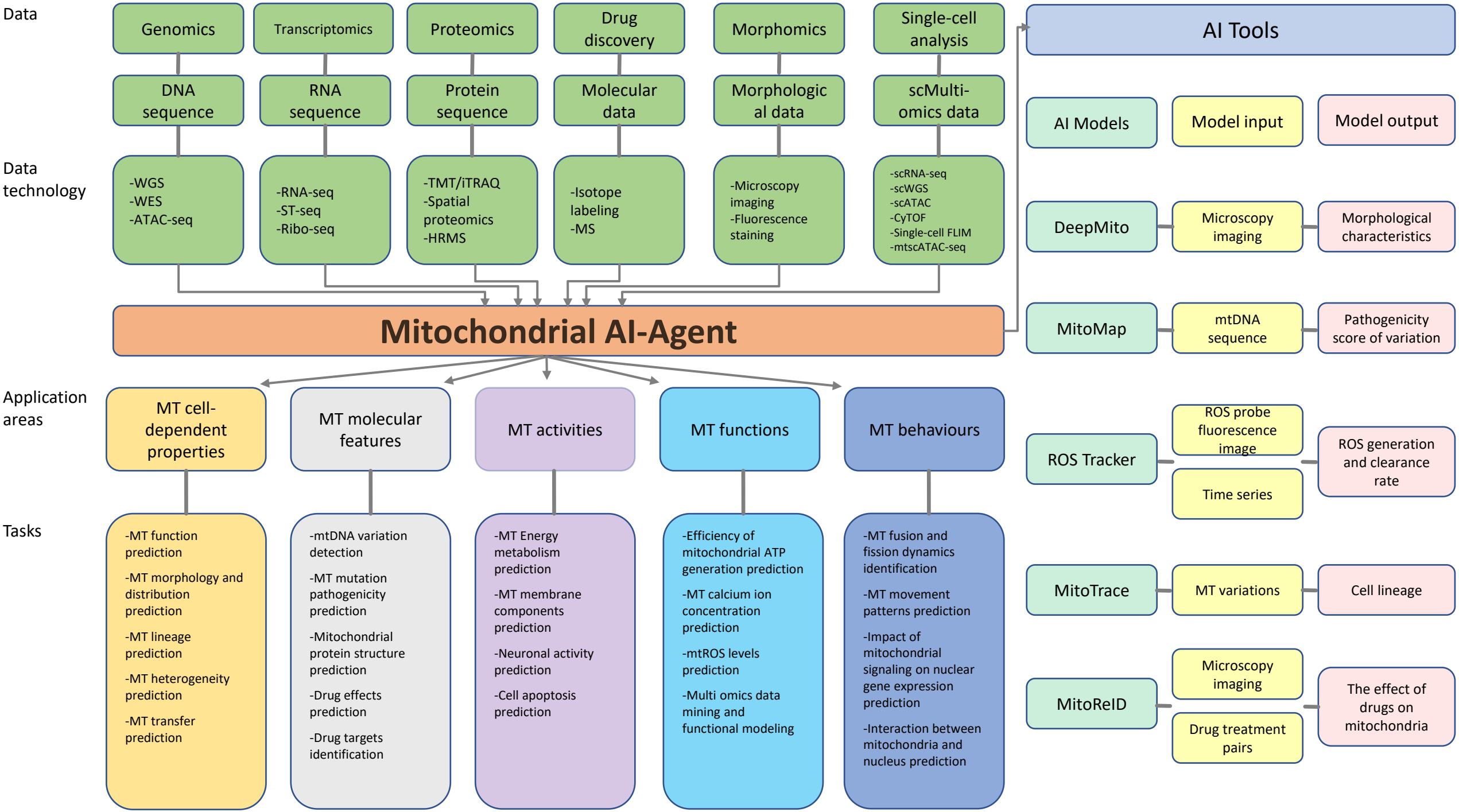
-Prediction of enzyme activity

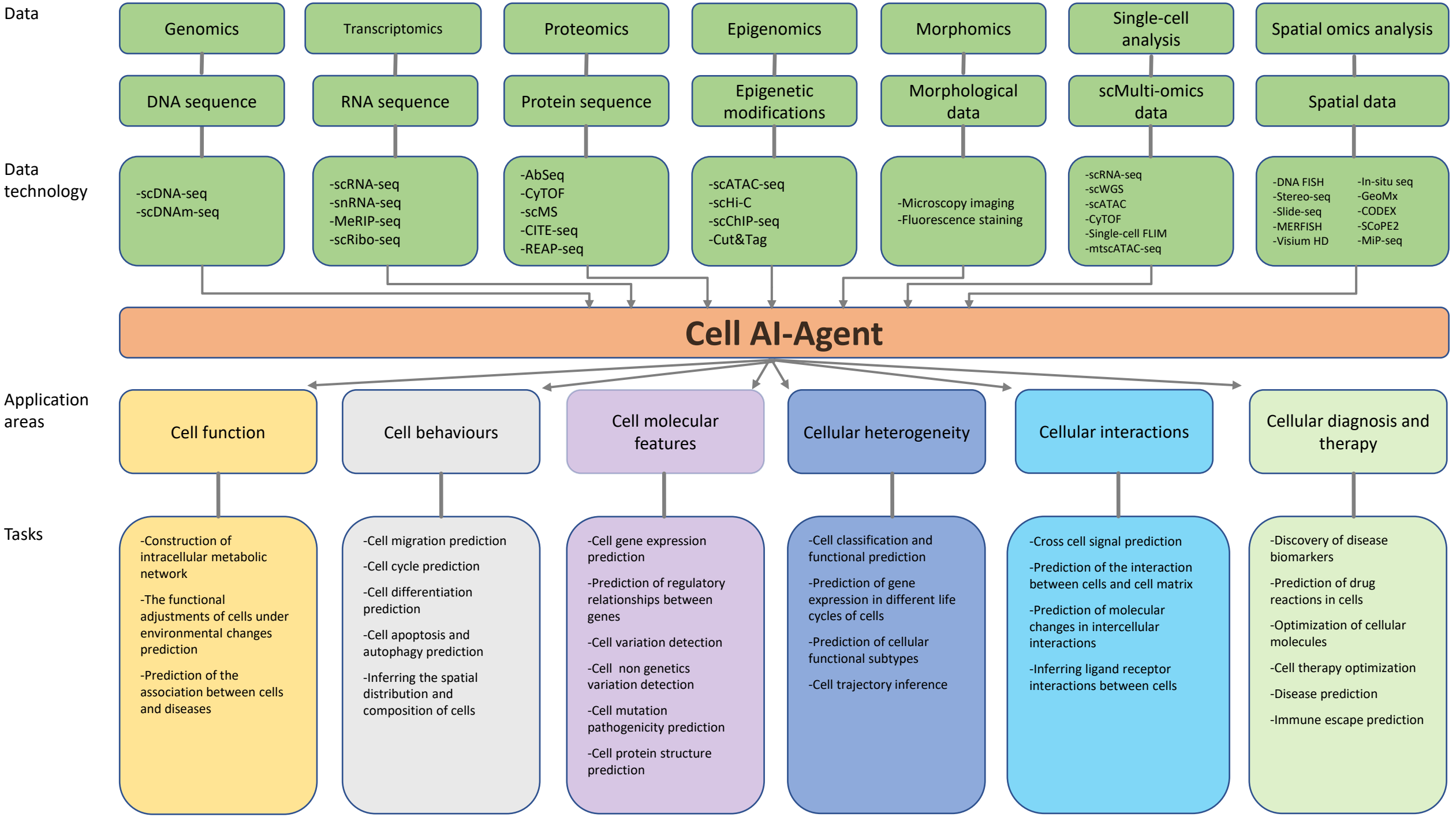
-Prediction of binding affinity and catalytic residues

-Prediction of changes in protein function caused by gene mutations

Drug design and engineering

-Protein-3D structure prediction
-Binding site prediction
-Guided molecule generation
-ADMET
-Docking score (Virtual screeing)
-MDS
-Similarity analysis
-Chemical synthesis
-Optimization of biochemical assay


Data
Genomics
Transcriptomics
Proteomics
Drug discovery
Morphomics
Single-cell analysis
DNA sequence
RNA sequence
Protein sequence
Molecular data
Morphological data
scMulti-omics data
Data technology
-WGS
-WES
-ATAC-seq
-RNA-seq
-ST-seq
-Ribo-seq
-TMT/iTRAQ
-Spatial proteomics
-HRMS
-Isotope labeling
-MS
-Microscopy imaging
-Fluorescence staining
-scRNA-seq
-scWGS
-scATAC
-CyTOF
-Single-cell FLIM
-mtscATAC-seq
Mitochondrial AI-Agent
Application areas
MT cell-dependent properties
MT molecular features
MT activities
MT functions
MT behaviours
Tasks
-MT function prediction
-MT morphology and distribution prediction
-MT lineage prediction
-MT heterogeneity prediction
-MT transfer prediction
-mtDNA variation detection
-MT mutation pathogenicity prediction
-Mitochondrial protein structure prediction
-Drug effects prediction
-Drug targets identification
-MT Energy metabolism prediction
-MT membrane components prediction
-Neuronal activity prediction
-Cell apoptosis prediction
-Efficiency of mitochondrial ATP generation prediction
-MT calcium ion concentration prediction
-mtROS levels prediction
-Multi omics data mining and functional modeling
-MT fusion and fission dynamics identification
-MT movement patterns prediction
-Impact of mitochondrial signaling on nuclear gene expression prediction
-Interaction between mitochondria and nucleus prediction
AI Tools
AI Models
Model input
Model output
DeepMito
Microscopy imaging
Morphological characteristics
MitoMap
mtDNA sequence
Pathogenicity score of variation
ROS Tracker
ROS probe fluorescence image
Time series
ROS generation and clearance rate
MitoTrace
MT variations
Cell lineage
MitoReID
Microscopy imaging
Drug treatment pairs
The effect of drugs on mitochondria


Data
Genomics
Transcriptomics
Proteomics
Epigenomics
Morphomics
Single-cell analysis
Spatial omics analysis
DNA sequence
RNA sequence
Protein sequence
Epigenetic modifications
Morphological data
scMulti-omics data
Spatial data
Data technology
-scDNA-seq
-scDNAm-seq
-scRNA-seq
-snRNA-seq
-MeRIP-seq
-scRibo-seq
-AbSeq
-CyTOF
-scMS
-CITE-seq
-REAP-seq
-scATAC-seq
-scHi-C
-scChIP-seq
-Cut&Tag
-Microscopy imaging
-Fluorescence staining
-scRNA-seq
-scWGS
-scATAC
-CyTOF
-Single-cell FLIM
-mtscATAC-seq
-DNA FISH
-Stereo-seq
-Slide-seq
-MERFISH
-Visium HD
-In-situ seq
-GeoMx
-CODEX
-SCoPE2
-MiP-seq
Cell AI-Agent
Application areas
Cell function
Cell behaviours
Cell molecular features
Cellular heterogeneity
Cellular interactions
Cellular diagnosis and therapy
Tasks
-Construction of intracellular metabolic network
-The functional adjustments of cells under environmental changes prediction
-Prediction of the association between cells and diseases
-Cell migration prediction
-Cell cycle prediction
-Cell differentiation prediction
-Cell apoptosis and autophagy prediction
-Inferring the spatial distribution and composition of cells
-Cell gene expression prediction
-Prediction of regulatory relationships between genes
-Cell variation detection
-Cell non genetics variation detection
-Cell mutation pathogenicity prediction
-Cell protein structure prediction
-Cell classification and functional prediction
-Prediction of gene expression in different life cycles of cells
-Prediction of cellular functional subtypes
-Cell trajectory inference
-Cross cell signal prediction
-Prediction of the interaction between cells and cell matrix
-Prediction of molecular changes in intercellular interactions
-Inferring ligand receptor interactions between cells
-Discovery of disease biomarkers
-Prediction of drug reactions in cells
-Optimization of cellular molecules
-Cell therapy optimization
-Disease prediction
-Immune escape prediction

Data

Application areas

Task refinement and execution

Genomics

-scDNA-seq
-scDNAm-seq

Transcriptomics

-scRNA-seq
-snRNA-seq
-MeRIP-seq
-scRibo-seq

Proteomics

-AbSeq
-CyTOF
-scMS
-CITE-seq
-REAP-seq

Epigenomics

-scATAC-seq
-scHi-C
-scChIP-seq
-Cut&Tag

Morphomics

-Microscopy imaging
-Fluorescence staining
-H&E imaging

Single-cell analysis

-scRNA-seq
-scWGS
-scATAC
-CyTOF
-scFLIM

Spatial omics analysis

-DNA FISH
-Stereo-seq
-Slide-seq
-MERFISH
-Visium HD

Tissue AI-Agent

Tissue function

-Simulate functional dynamics within an organization

-Predicting the impact of different environments on organizational function

Tissue behaviours

-Prediction of the dynamic behavior of an organization

-Prediction of cell migration, proliferation, and apoptosis within tissues

Tissue molecular features

-Identify tissue-specific molecular markers

-Prediction of molecular changes in organizations

Tissue heterogeneity

-Prediction of heterogeneity in different regions of an organization, including cell types, metabolic activity, and molecular distribution

-Prediction of the spatial stratification and distribution of organizational functions

Tissue diagnosis and therapy

-Prediction of drug response to specific tissues

-Optimizing tissue engineering and regenerative medicine solutions

Data

Genomics | Transcriptomics | Proteomics | Physiology | Morphomics | Single-cell analysis | Spatial omics analysis

DNA sequence | RNA sequence | Protein sequence | Physiological data | Morphological data | scMulti-omics data | Spatial data

Organ AI-Agent

Application areas

Organ function modeling | Organ pathophysiology analysis | Organ molecular features | Organ heterogeneity | Tissue interactions | Drug response and toxicity analysis

Task refinement and execution

- Organ function modeling:
  - -Simulate and predict the functional status of a single organ under healthy and diseased conditions
  - -Simulate and predict the impact of the environment on organ function
- Organ pathophysiology analysis:
  - -Dynamic changes of simulator specific diseases
  - -Prediction of the causes of organ dysfunction
- Organ molecular features:
  - -Identifying organ specific marker genes
  - -Molecular characteristics under pathological progression prediction
- Organ heterogeneity:
  - -Heterogeneity in different regions, including cell types, metabolic activity, and molecular distribution
  - -Prediction of the spatial stratification and distribution of organizational functions
- Tissue interactions:
  - -Building a communication network between organizations
  - -Simulate signal transmission between organizations
- Drug response and toxicity analysis:
  - -Prediction of the effects of drugs on specific organs
  - -Optimizing organ engineering and regenerative medicine solutions
  - -Simulating the adaptation process of transplanted organs in the host

Data

Application areas

Task refinement and execution

Whole-System Imaging data

-MRI
-CT
-PET

Physiological data

-Blood pressure
-Heart rate,
-Respiratory rate
-Drugs

Multi-Omics data

-GWAS
-RNA-seq
-WGS

Anatomy data

-Thickness
-Density,
-Mechanical properties

Electronic Medical Record (EMR)

-Population statistics;
-Diagnosis;
-Clinical records

Organ System AI-Agent

Multi-organ functional coordination

-Building functional related networks for multiple organs

-Prediction of how multiple organ dysfunction can lead to systemic diseases

Signal transmission

-Prediction of the propagation path of cellular signals between multiple organs

-Analyzing how signal imbalance between organs triggers pathological cascade effects

Multi-organ pathological cascades

-Simulate the dynamic process of disease spreading from one organ to other organs

Multi-organ damage and repair

-Simulating the process of cross organ compensation and repair after organ injury

Multi-organ metabolic networks

-Constructing a metabolic flow and balance network among multiple organs

-Prediction of the flow and changes of metabolites in multiple organs

Multi-organ biomechanical interactions

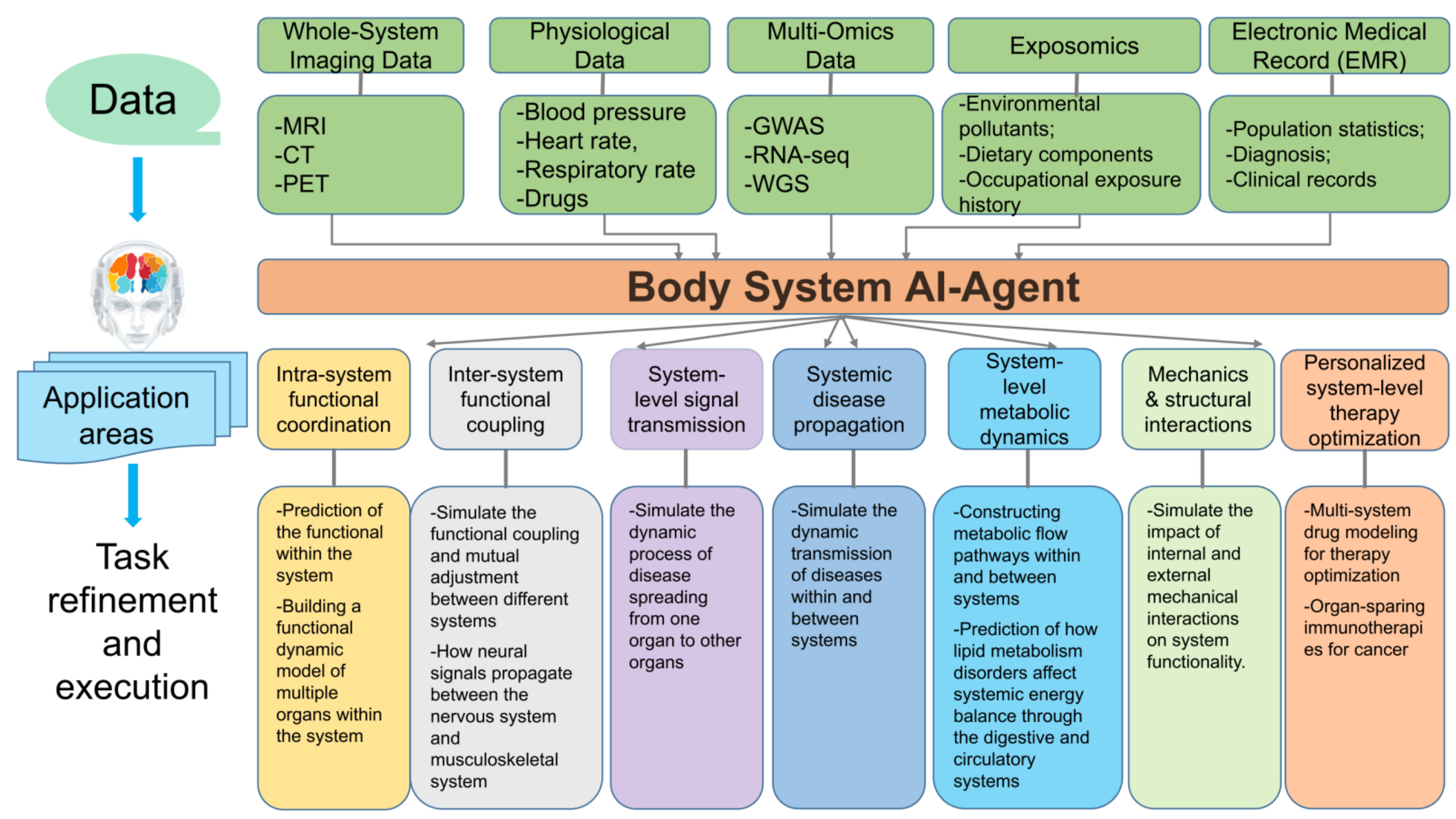
-Analyze the impact of inter organ interactions on functional performance

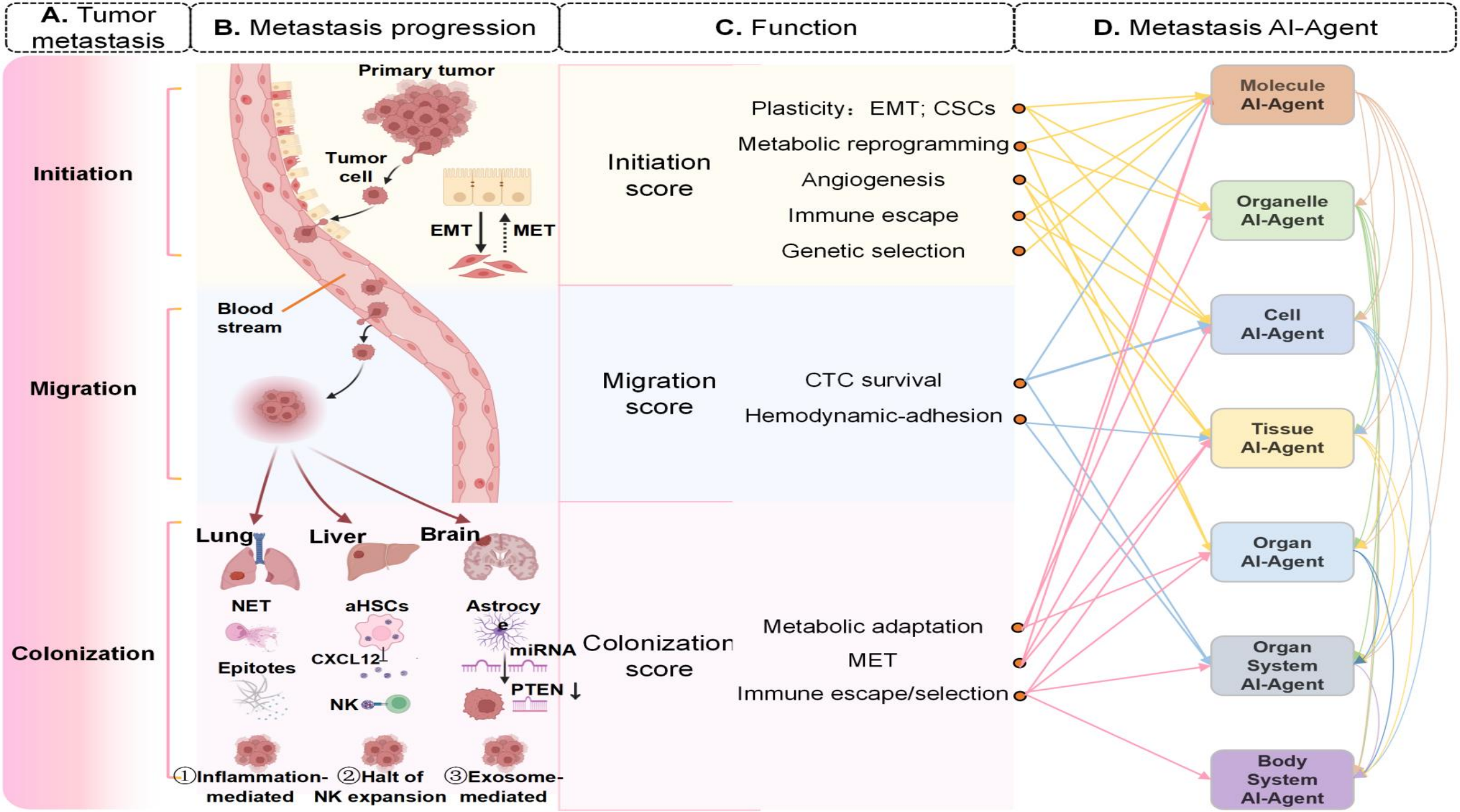
-Prediction of phenotype related organs and organ changes

Multi-organ responses of drugs

-Optimize multi organ treatment plan

-Prediction of the multi organ combined effects of drugs


Data
Whole-System Imaging Data
-MRI
-CT
-PET
Physiological Data
-Blood pressure
-Heart rate,
-Respiratory rate
-Drugs
Multi-Omics Data
-GWAS
-RNA-seq
-WGS
Exposomics
-Environmental pollutants;
-Dietary components
-Occupational exposure history
Electronic Medical Record (EMR)
-Population statistics;
-Diagnosis;
-Clinical records
Body System AI-Agent
Application areas
Intra-system functional coordination
-Prediction of the functional within the system
-Building a functional dynamic model of multiple organs within the system
Inter-system functional coupling
-Simulate the functional coupling and mutual adjustment between different systems
-How neural signals propagate between the nervous system and musculoskeletal system
System-level signal transmission
-Simulate the dynamic process of disease spreading from one organ to other organs
Systemic disease propagation
-Simulate the dynamic transmission of diseases within and between systems
System-level metabolic dynamics
-Constructing metabolic flow pathways within and between systems
-Prediction of how lipid metabolism disorders affect systemic energy balance through the digestive and circulatory systems
Mechanics & structural interactions
-Simulate the impact of internal and external mechanical interactions on system functionality.
Personalized system-level therapy optimization
-Multi-system drug modeling for therapy optimization
-Organ-sparing immunotherapies for cancer
Task refinement and execution


A. Tumor metastasis
B. Metastasis progression
C. Function
D. Metastasis AI-Agent
Initiation
Migration
Colonization
Primary tumor
Tumor cell
EMT
MET
Blood stream
Lung
Liver
Brain
NET
aHSCs
Astrocyte
Epitotes
CXCL12
miRNA
PTEN
NK
①Inflammation-mediated
②Halt of NK expansion
③Exosome-mediated
Initiation score
Plasticity： EMT; CSCs
Metabolic reprogramming
Angiogenesis
Immune escape
Genetic selection
Migration score
CTC survival
Hemodynamic-adhesion
Colonization score
Metabolic adaptation
MET
Immune escape/selection
Molecule AI-Agent
Organelle AI-Agent
Cell AI-Agent
Tissue AI-Agent
Organ AI-Agent
Organ System AI-Agent
Body System AI-Agent

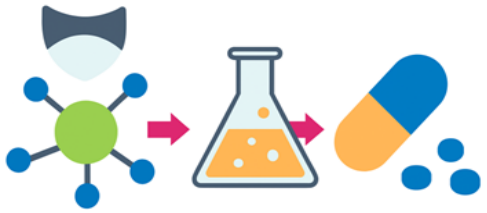
AI Generative Models for Drug Design

Pipeline of drug development	Target identification	Drug discovery	Preclinical studies (animal)	Clinical trials (human)	Approval/launch	Post-market surveillance
Main tasks	New target 	Discovery of active compounds (hits/leads) Lead optimization for drug candidates CMC	Biomarker discovery DMPK evaluation Safety evaluation Pharmacodynamics evaluation Medication scheme		New drug 	Evaluating or monitoring - Therapeutic efficacy - Side effect - Drug stability - Medication plan Failure analysis Counterfeit analysis
AI-powered applications	• Multi-omics data analysis • Biological network construction and analysis • Literature and real-word data mining • Knowledge graph construction • Target validation	• Virtual screening • Prediction of ligand–receptor interaction • Molecular generation • ADMET prediction • Lead optimization • Synthesis route planning • Automated synthesis	• Diagnostic, predictive and prognostic biomarkers • Pharmacometrics property prediction • Clinical trial success prediction • Prediction of drug side effects • Digital Twins in clinical trial design • Drug repurposing • Regulatory approval support		• Personalized efficacy assessment • Personalizing patient care • Early detection of safety issues • Automated adverse event reporting • Continuous safety monitoring • Drug benefit–risk assessment • Enhancing pharmacovigilance compliance	

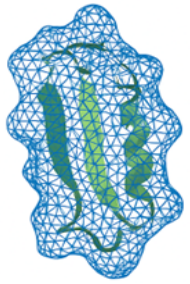
Zhang, et. Al., AI in Drug Development, Nature Medicine, 2025

Drug Designing

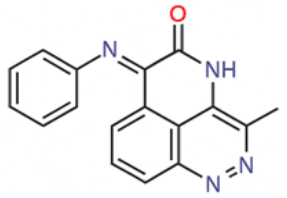
Denovo drug designing



GALILEO (2023), GCDM(2023)
GeoLDM(2023), JODO(2023)
MiDi(2023), MDM(2022)
RJT-RL (2022), MolPal (2021)
STONED (2021), E-NF (2021)
EDM(2022), CReM (2020)
DeLinker (2020)

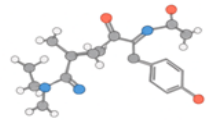


Target aware drug designing



TemGen(2024)
DiffDock (2023)
DrugGPT(2023)
TargetDiff(2023)
DiffSBDD(2022)
Pocket2Mol(2022)
LiGAN(2020)

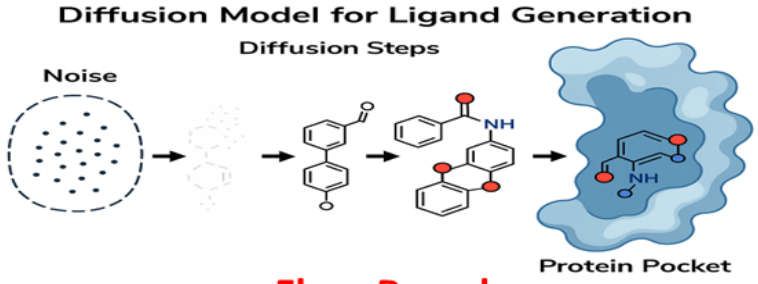
Conformation drug designing



GeoDiff(2022)
CGCF (2021)
GeoMol(2021)
CongGF(2021)
DGSM(2021)

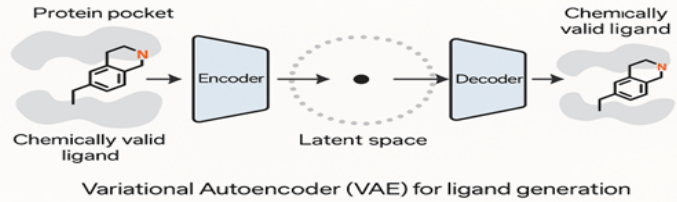
Generative AI model

Diffusion model



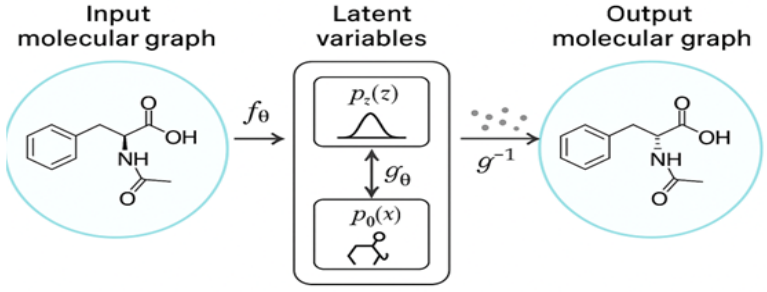
DiffDock
TorchDrug
DiffSBDD
DeepICL
DiffSMol
GeoDiff / EDM

Variational Autoencoder



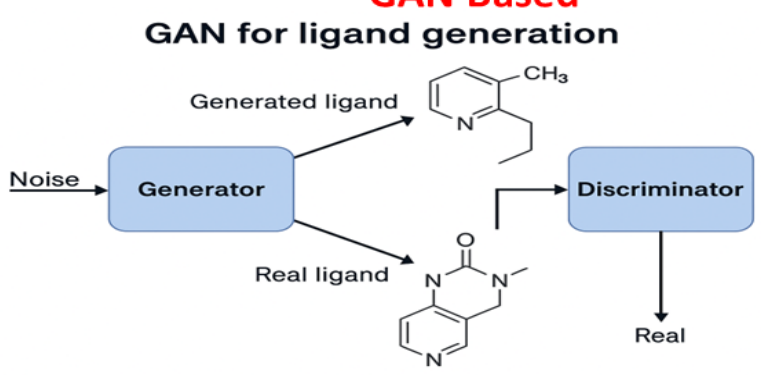
JT-VAE
Grammar VAE
GraphVAE
HierVAE
CVAE

Flow Based

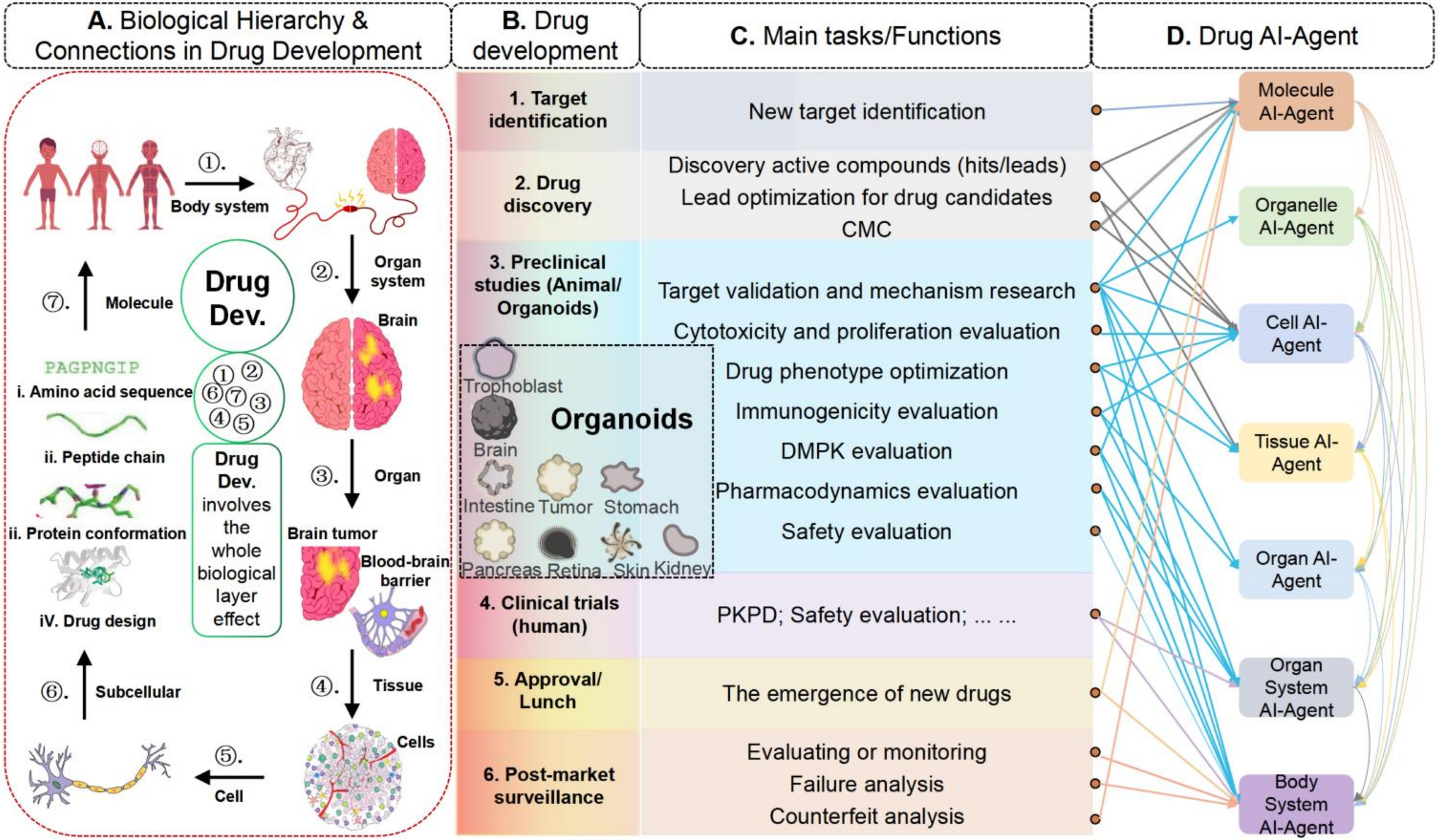


3DFlowMolecule
MoFlow
GraphAF
FlowMol
GeoMol
DiffFlow
GraphNVP
Invertible Attention Flow

GAN Based



MolGAN
ORGAN
ChemGAN
LatentGAN
GAN-ZINC
GraphGAN
ReLeaSE
Mol-CycleGAN


A. Biological Hierarchy & Connections in Drug Development
B. Drug development
C. Main tasks/Functions
D. Drug AI-Agent
Body system
Organ system
Brain
Organ
Brain tumor
Blood-brain barrier
Tissue
Cells
Cell
Subcellular
Molecule
Drug Dev.
PAGPNGIP
i. Amino acid sequence
ii. Peptide chain
ii. Protein conformation
iV. Drug design
Drug Dev. involves the whole biological layer effect
1. Target identification
New target identification
2. Drug discovery
Discovery active compounds (hits/leads)
Lead optimization for drug candidates
CMC
3. Preclinical studies (Animal/ Organoids)
Target validation and mechanism research
Cytotoxicity and proliferation evaluation
Drug phenotype optimization
Immunogenicity evaluation
DMPK evaluation
Pharmacodynamics evaluation
Safety evaluation
Organoids
Trophoblast
Brain
Intestine
Tumor
Stomach
Pancreas
Retina
Skin
Kidney
4. Clinical trials (human)
PKPD; Safety evaluation; ... ...
5. Approval/ Lunch
The emergence of new drugs
6. Post-market surveillance
Evaluating or monitoring
Failure analysis
Counterfeit analysis
Molecule AI-Agent
Organelle AI-Agent
Cell AI-Agent
Tissue AI-Agent
Organ AI-Agent
Organ System AI-Agent
Body System AI-Agent

Brief Summary

- Proposed the Full-Body AI-Agent framework, a multi-agent architecture designed to model human biology from molecular to whole-organism scales.
- Unlike existing biomedical AI systems confined to discrete tasks or domains, the framework integrates seven biologically grounded agents under a central coordination layer, enabling iterative, bidirectional reasoning across scales.
- By unifying multi-omics, imaging, physiological, and clinical data, the system constructs dynamic, system-wide mechanistic models that bridge molecular discovery with systemic simulation.
- Demonstrated in applications such as systemic disease modeling and drug development, the approach offers a coherent computational paradigm to reduce translational gaps, enhance predictive accuracy, and accelerate the development of safe and effective therapies.

Acknowledgements

- Research Collaborations:
  - Research Fellows: Jianguo Wen, Kexin Huang, Aoqi Wang, Tiangang Wang, Ruihan Luo, Lei You, Yanfei Wang, Jiajia Liu, Bongrui Li, Yujun Liu, Xiaona Liu, Zhiwei Fan
  - Other major collaborators: Drs. Bishehsari, Pora Kim, Hongyu Wang; Qianqian Song, Eduardo Matta, Qilin Cao, Lei Xie, Jake Chen, Claudio Soto

- Major Funding Resources

  - **NIH R01LM010185**
  - **NIH U01HL111560**
  - **NIH R01DE012074**
  - **NIH U54CA149196**
  - **NIH U01CA166886**
  - **NIH R01CA155069**
  - **NIH R01DE021863**
  - **NSF 2217515 and 2326879**

THANK YOU !

UTHealth Houston
McWilliams School of Biomedical Informatics

Xiaobo Zhou Ph.D.
FAIMBE, FACMI
Carl V Vartian Professor and Director,
McWilliams School of Biomedical Informatics,
University of Texas Health Science, Houston, TX