University of Connecticut
Economic Development Forum 2025

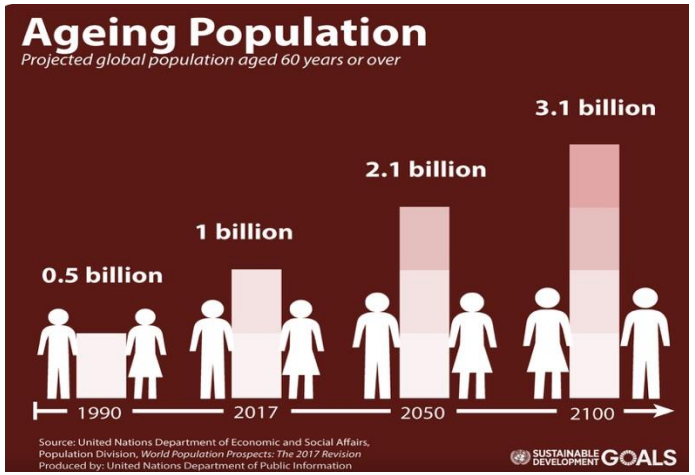
NEC Laboratories America

Agentic Generative AI for Precision Medicine

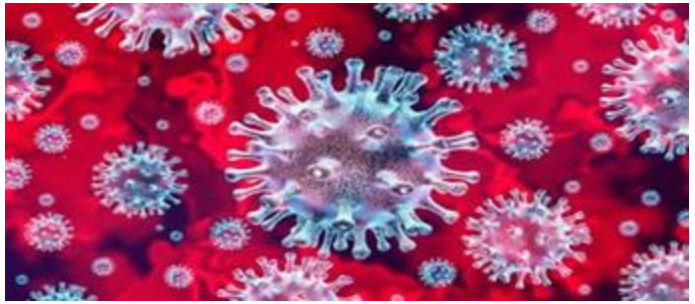
Martin Renqiang Min
NEC Laboratories America
Nov 13, 2025

Modern Healthcare Needs Transformation

A growing elderly population, driven by increases in the expected human lifespan, will pressure the existing medical infrastructure.



Widespread and low-cost world travel facilitates the quickly and easily spread of disease between countries, increasing the potential of frequent global pandemics.



Medical knowledge explosion will require an ever-increasing number of medical experts to manage and deliver life saving treatments.

Traditional Drug Design Business is Not Sustainable

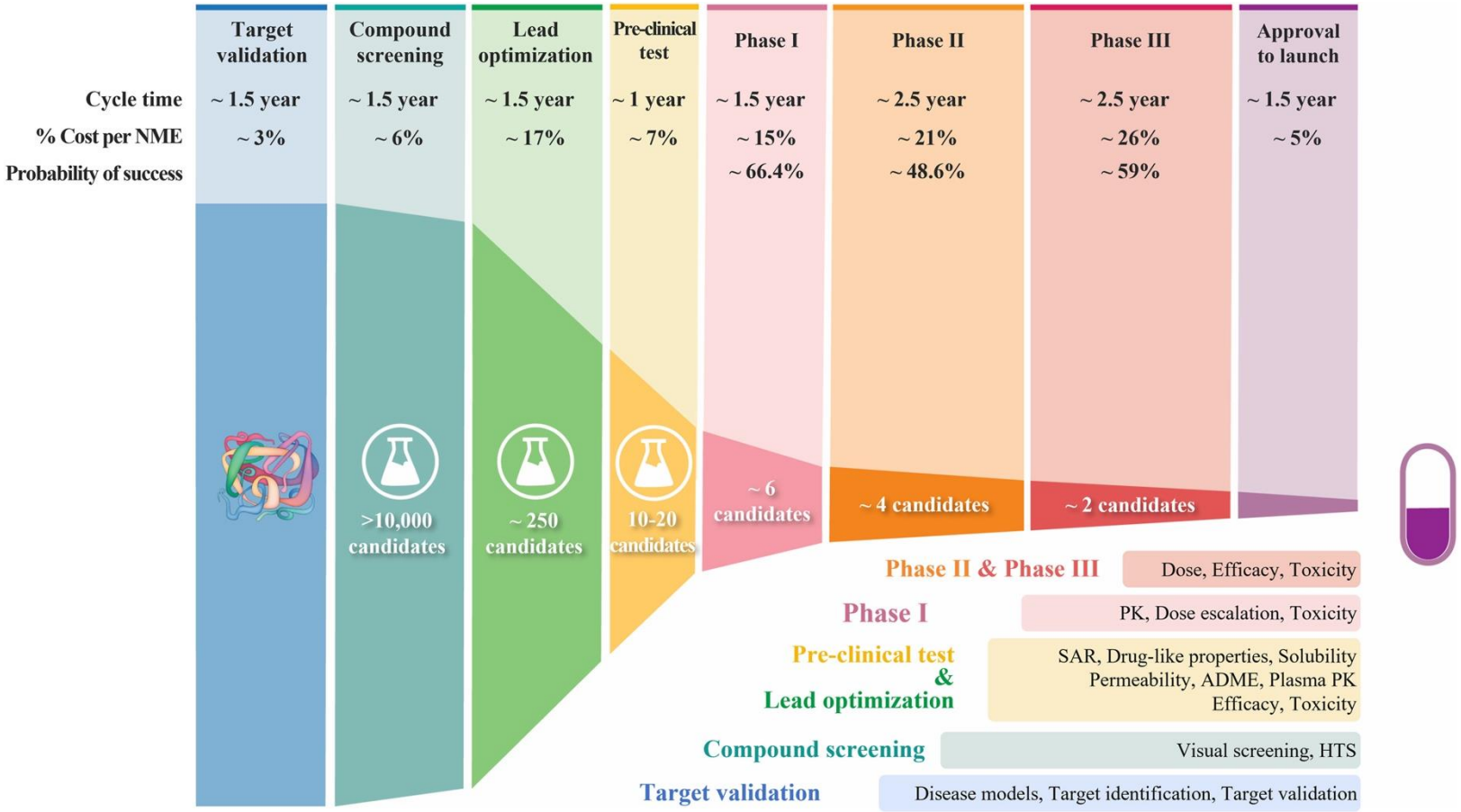


Figure resource: theconversation.com

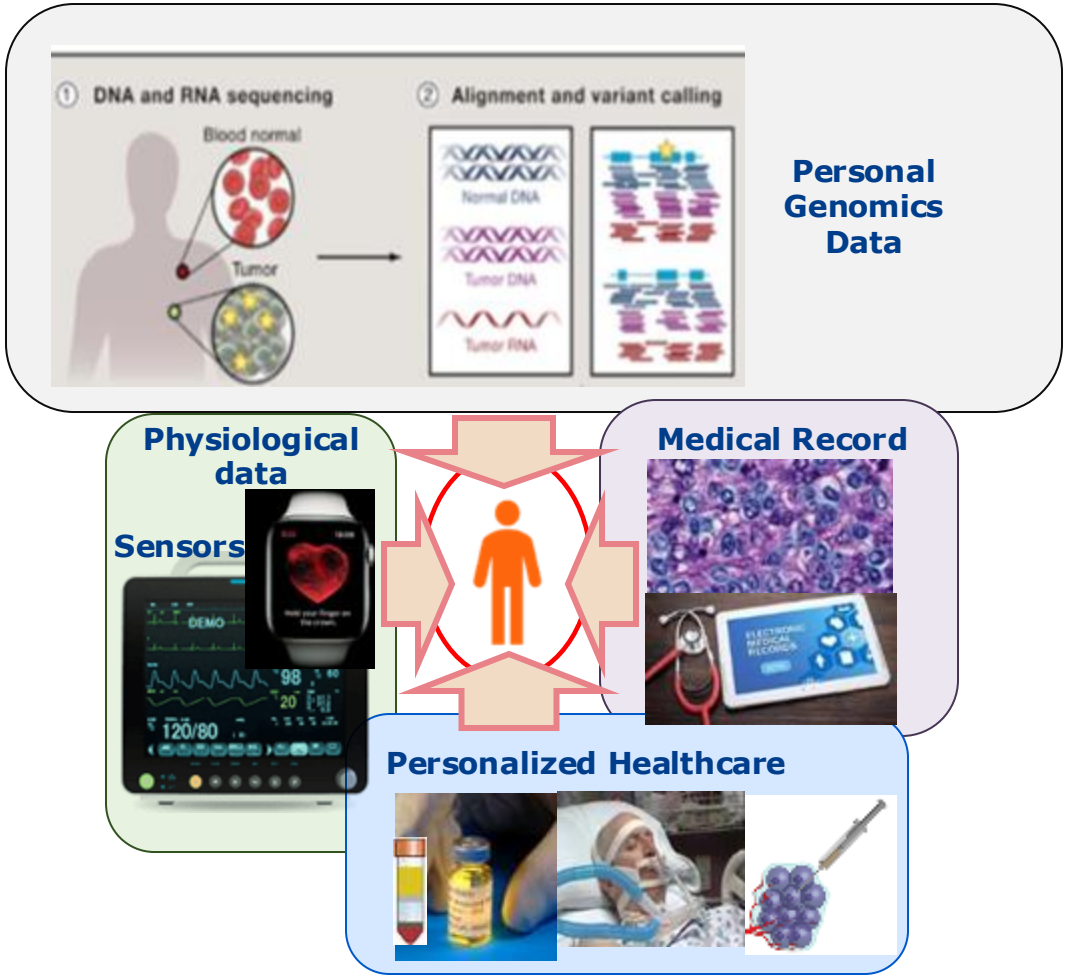
Increasing pharmaceutical R&D cost, but still **low success rate** (<10%)

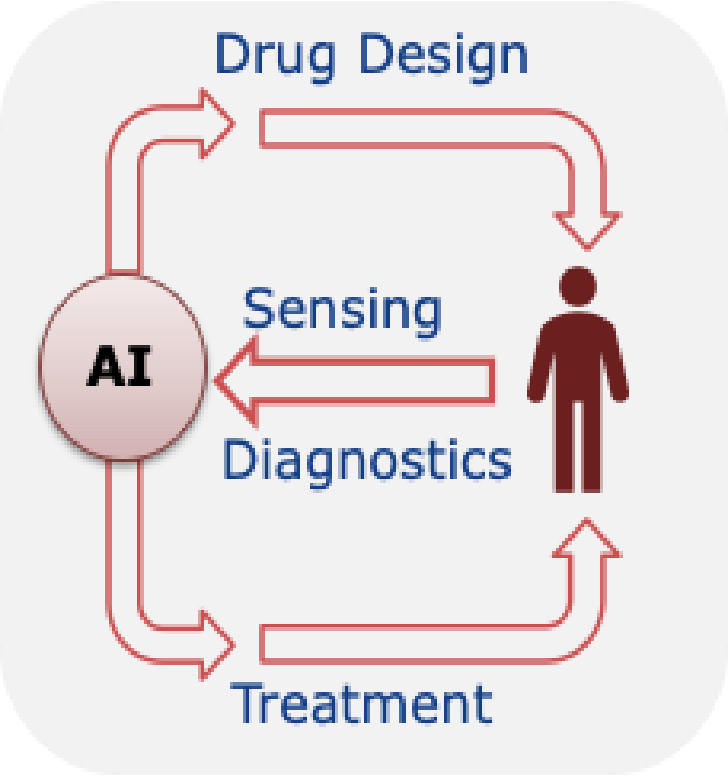
Average time/cost for designing one new drug = >10 years + $~2B

Large-Scale Biomedical Data Enables Precision Medicine and Healthcare

TCGA's Pan-Cancer Atlas

TCGA Outcomes & Impact





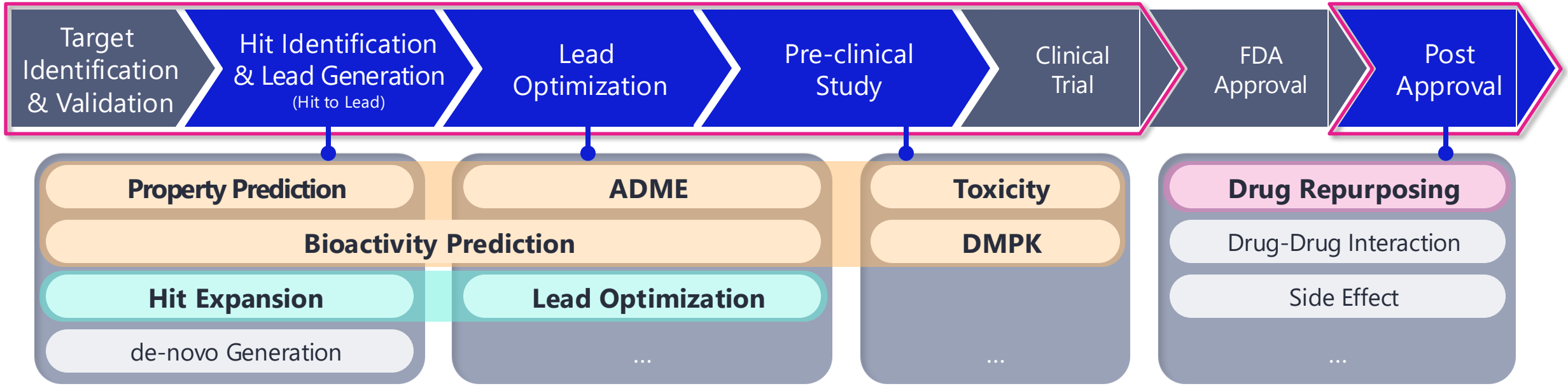
AI Accelerated Drug Development Process

Advances in **AI and high-throughput experiments** empower **accelerated drug development**

FDA NEWS RELEASE

FDA Announces Plan to Phase Out Animal Testing Requirement for Monoclonal Antibodies and Other Drugs

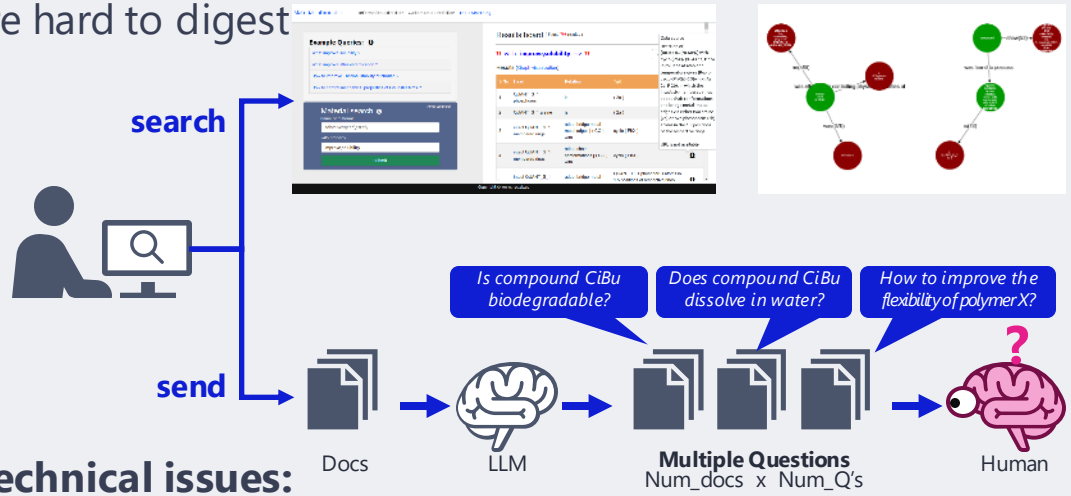
For Immediate Release: April 10, 2025



Human Collaborative Multi-Agents for Drug Screening

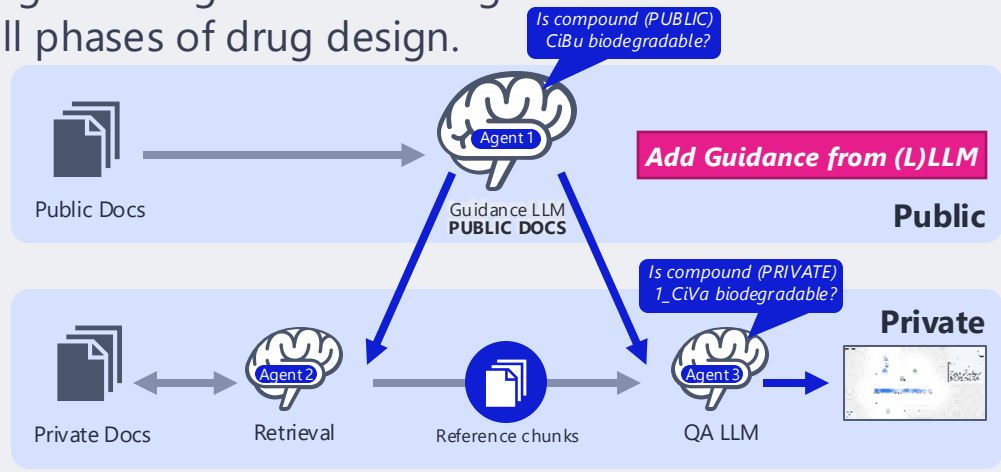
Enable interactive privacy-preserving insight discovery from multimodal data to significantly reduce drug investigation, screening, and design time

Traditional: Scientists manually search thousands of research papers to investigate/screen drug candidates or send a lot of private data to public LLMs to get lengthy textual answers that are hard to digest



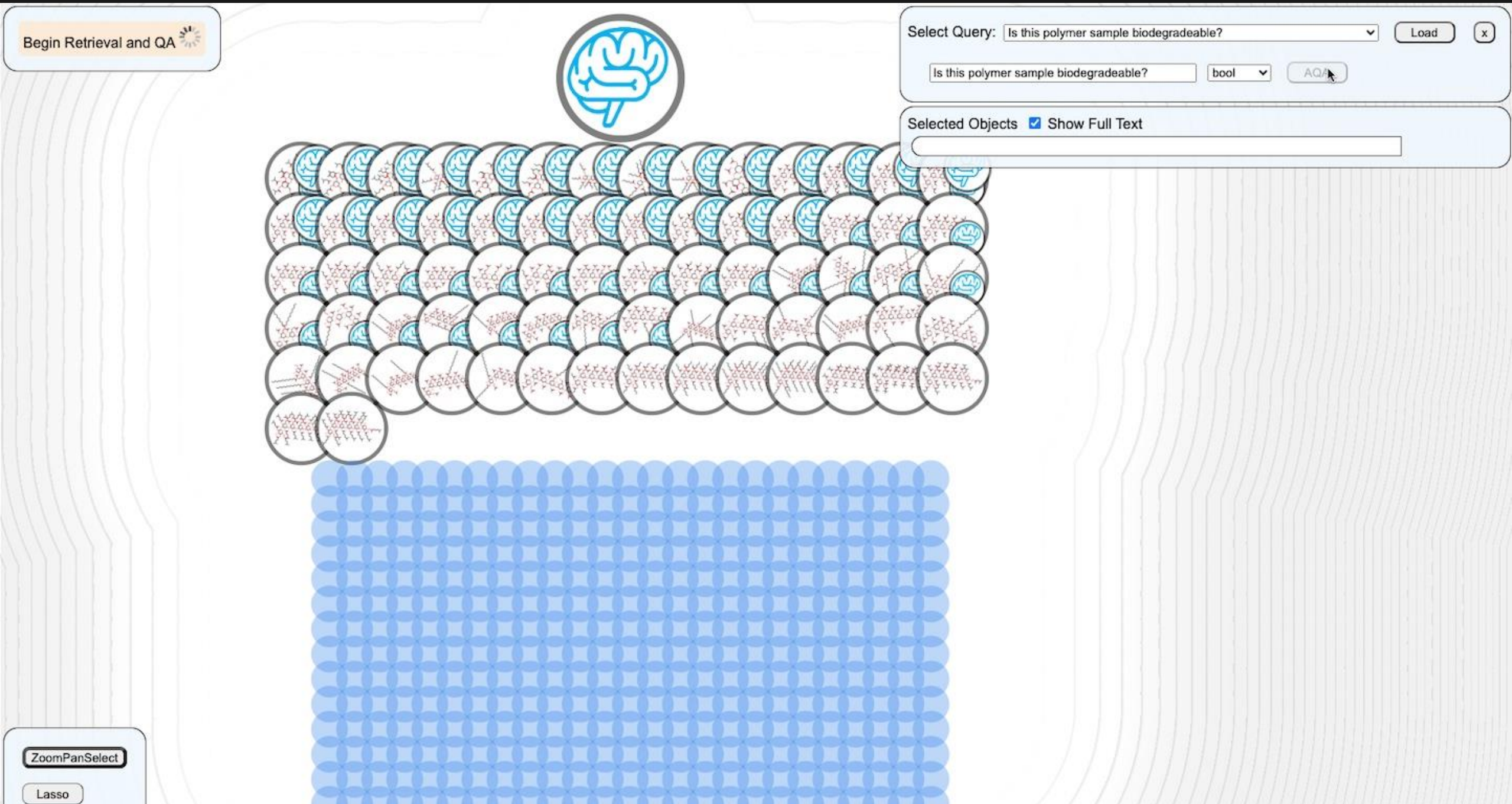
Technical issues:
Reading many literature papers and reports is time-consuming; Sending private company data to public LLMs poses data-leaking and privacy issues.

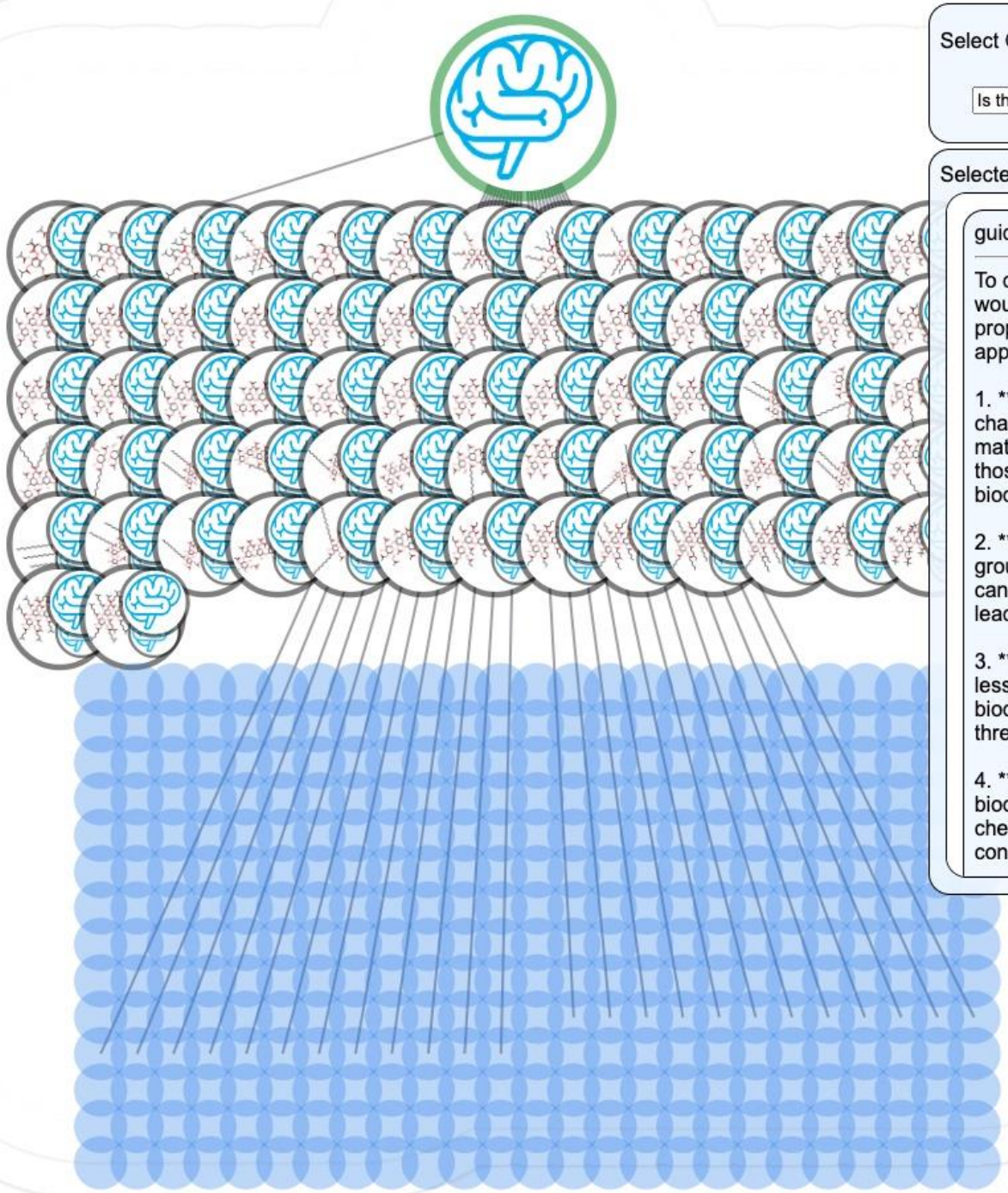
NEC: Human collaborative AI agents for interactively visualizing large-scale existing drugs, filtering candidates, answering questions, and generating manufacturing methods to increase efficiency in all phases of drug design.



Efficient and safe drug insight discovery:
LLM-enhanced interactive data visualization and QA increases the efficiency of insight discovery for drug design; Local agents enable data safety.

Obuchi *et al.*, LLMs and MI Bring Innovation to Material Development Platforms. NEC Technical Journal, Special Issue on Revolutionizing Business Practices with Generative AI. 2024.


Begin Retrieval and QA
Select Query:
Is this polymer sample biodegradeable?
Load
x
Is this polymer sample biodegradeable?
bool
AQA
Selected Objects
Show Full Text
ZoomPanSelect
Lasso



Select Query: Is this polymer sample biodegradeable? Load x

Is this polymer sample biodegradeable? bool AQA

Selected Objects Show Full Text

guidance_0 x

To determine if the provided polymer sample is biodegradable, I would analyze the information given and cross-reference it with the properties known to influence biodegradability. Here's how I would approach it:

1. **Chemical Structure**: I would examine the main chain and side chain composition to see if they consist of naturally occurring materials (like cellulose) or synthetic materials. Natural polymers or those derived from natural materials are more likely to be biodegradable.

2. **Functional Groups**: I would look into the type of functional groups present in the polymer structure. Ester groups, for example, can be subject to hydrolysis under the right conditions, which can lead to degradation.

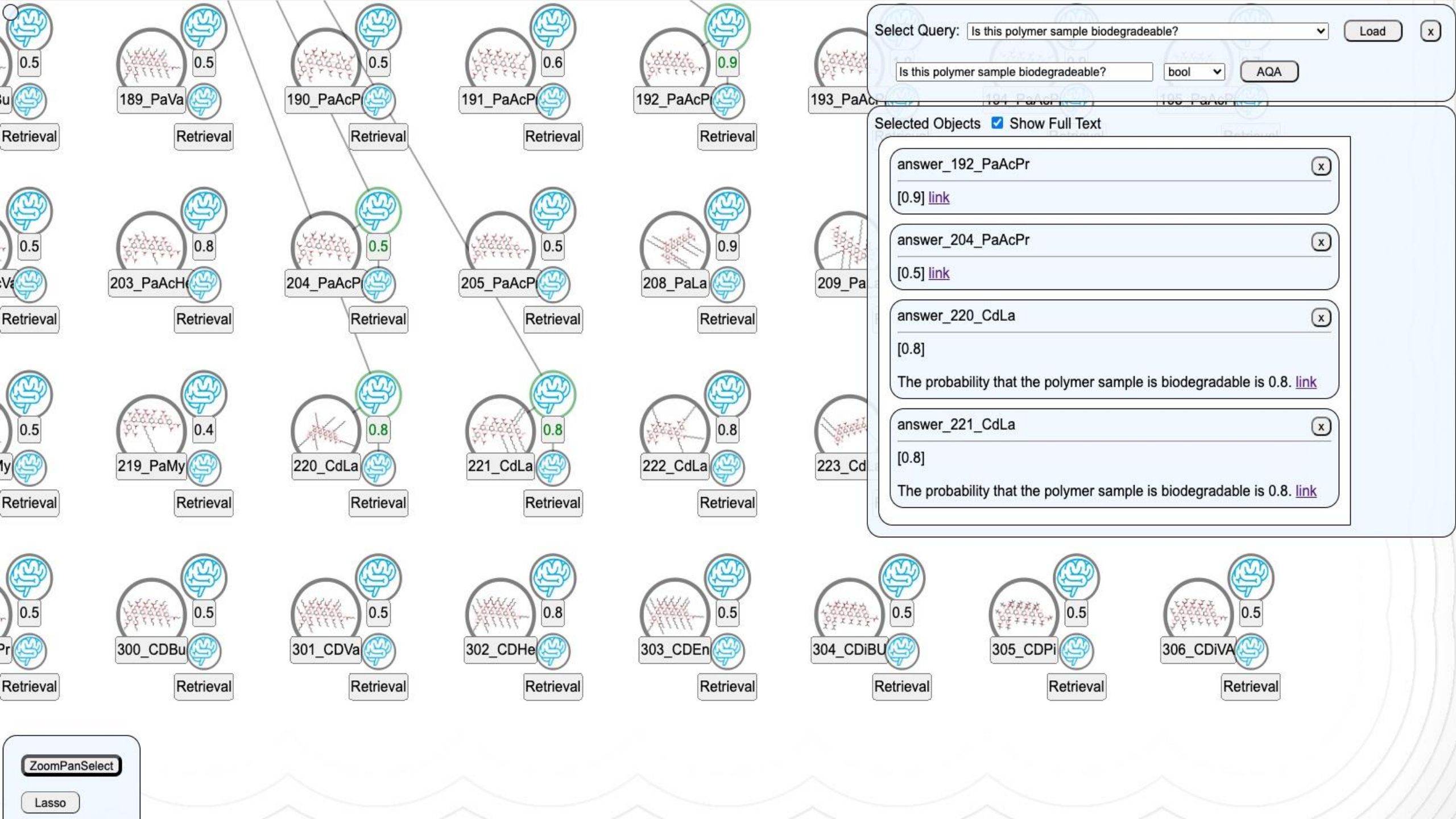
3. **Degree of Substitution (DS)**: The higher the DS, typically, the less soluble the polymer might be, which might impact biodegradability. I would compare the DS value provided with known thresholds that might predict biodegradability.

4. **Molecular Weight**: Higher molecular weights can hinder biodegradation; however, there can be exceptions depending on the chemical structure. The molecular weight provided would be considered in the context of known data for similar polymers.

ZoomPanSelect

Lasso

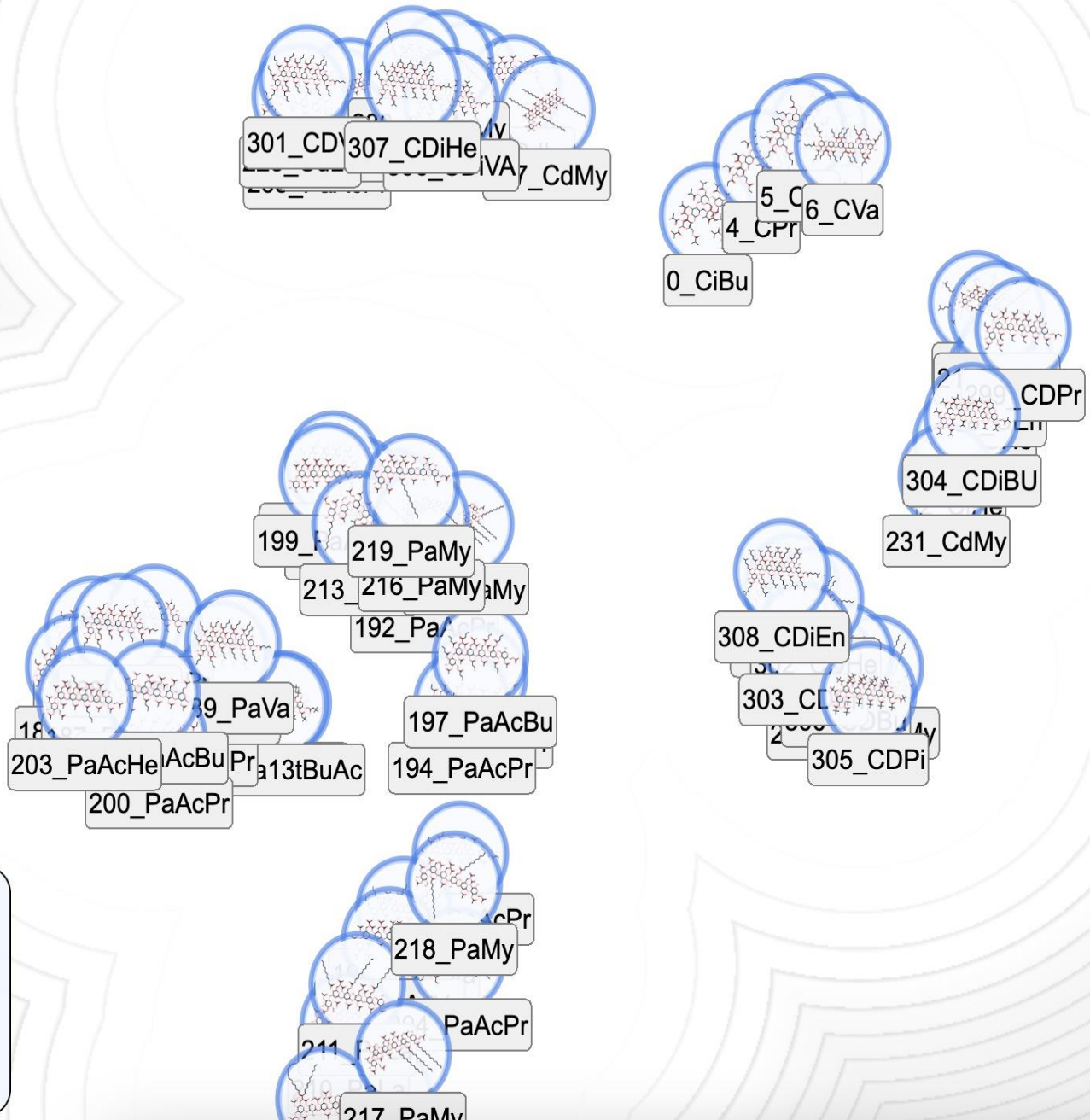
Zoom Fit


Select Query: Is this polymer sample biodegradeable?
Load
x
Is this polymer sample biodegradeable?
bool
AQA
Selected Objects
Show Full Text
answer_192_PaAcPr
[0.9] link
answer_204_PaAcPr
[0.5] link
answer_220_CdLa
[0.8]
The probability that the polymer sample is biodegradable is 0.8. link
answer_221_CdLa
[0.8]
The probability that the polymer sample is biodegradable is 0.8. link
189_PaVa
190_PaAcP
191_PaAcP
192_PaAcP
193_PaAc
203_PaAcHe
204_PaAcP
205_PaAcP
208_PaLa
209_Pa
219_PaMy
220_CdLa
221_CdLa
222_CdLa
223_Cd
300_CDBu
301_CDVa
302_CDHe
303_CDEn
304_CDiBU
305_CDPi
306_CDiVA
Retrieval
ZoomPanSelect
Lasso

1.	Is this polymer sample biodegradeable?	bool	☑	embed
2.	Is this polymer sticky?	bool	☑	embed
3.	Is this polymer sample flexible?	bool	☑	embed
4.	Is this polymer sample based on curdlan?	bool	☑	embed
5.	What is the molecular weight of this polymer s	numeric	☑	embed
6.	Will this polymer sample float on water?	bool	☑	embed
7.	Is this polymer sample going to save the world	bool	☑	embed
8.	What is the density of this polymer sample?	numeric	☑	embed

Render

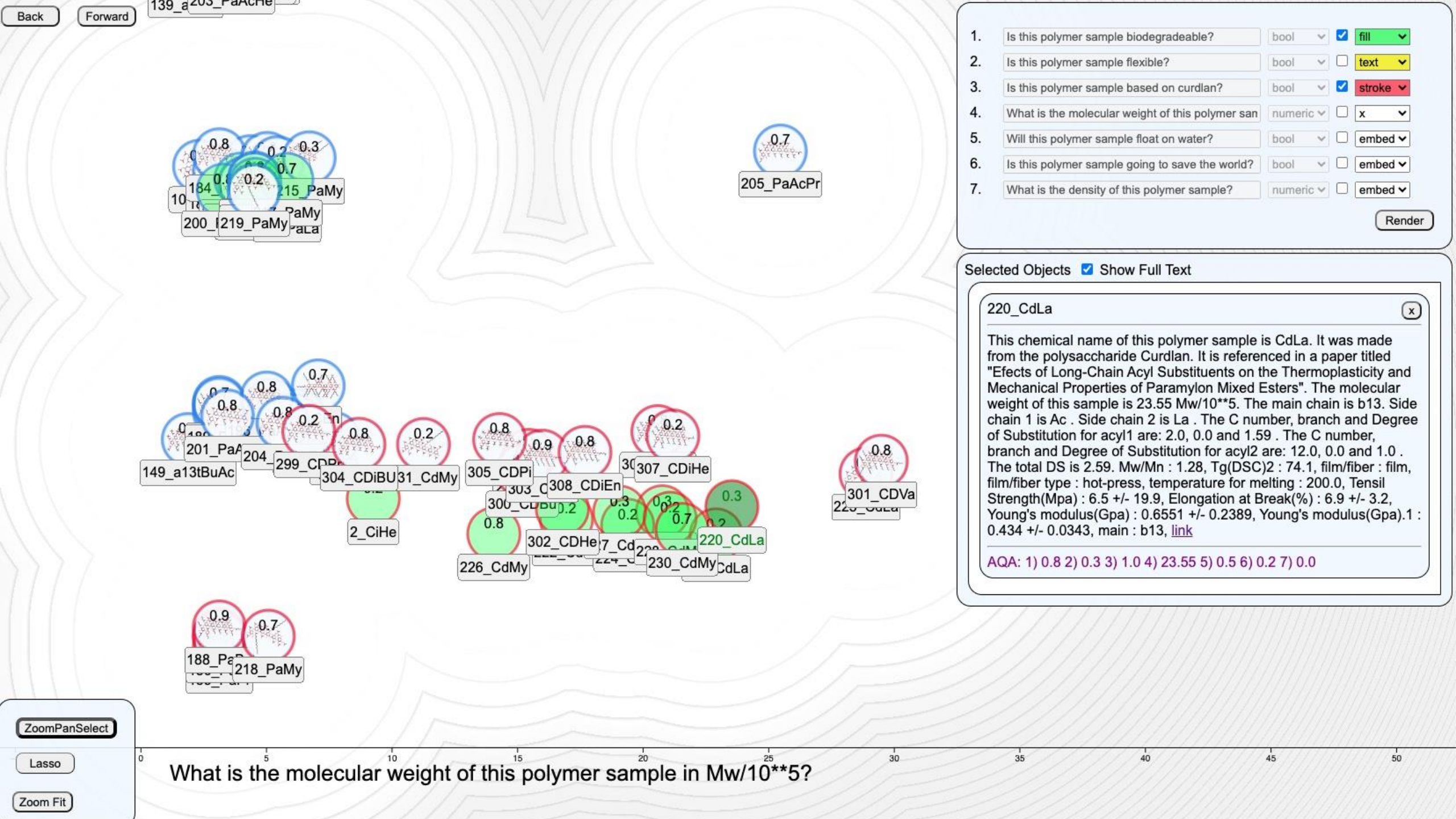
Selected Objects ☑ Show Full Text



1.	Is this polymer sample biodegradeable?	bool	☑	embed
2.	Is this polymer sticky?	bool	☑	embed
3.	Is this polymer sample flexible?	bool	☑	embed
4.	Is this polymer sample based on curdlan?	bool	☑	embed
5.	What is the molecular weight of this polymer s	numeric	☑	embed
6.	Will this polymer sample float on water?	bool	☑	embed
7.	Is this polymer sample going to save the world	bool	☑	embed
8.	What is the density of this polymer sample?	numeric	☑	embed

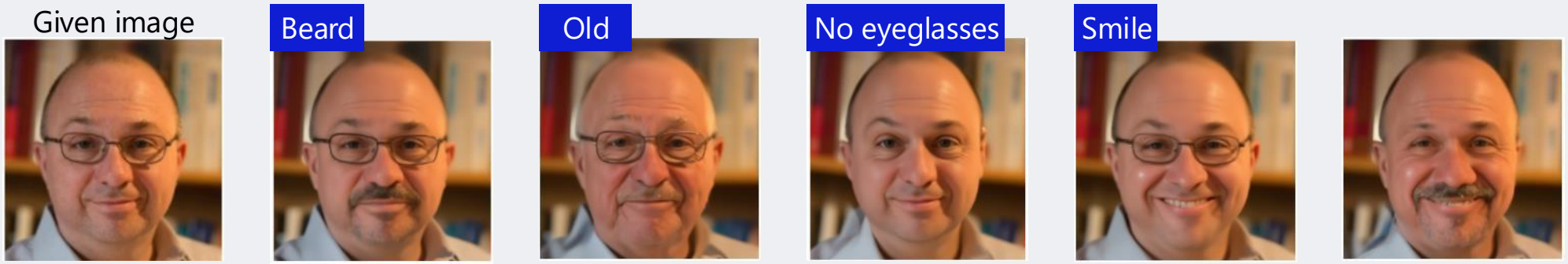
Render

Selected Objects ☑ Show Full Text


Back
Forward
1. Is this polymer sample biodegradeable? bool fill
2. Is this polymer sample flexible? bool text
3. Is this polymer sample based on curdlan? bool stroke
4. What is the molecular weight of this polymer san numeric x
5. Will this polymer sample float on water? bool embed
6. Is this polymer sample going to save the world? bool embed
7. What is the density of this polymer sample? numeric embed
Render
Selected Objects
Show Full Text
220_CdLa
This chemical name of this polymer sample is CdLa. It was made from the polysaccharide Curdlan. It is referenced in a paper titled "Efects of Long-Chain Acyl Substituents on the Thermoplasticity and Mechanical Properties of Paramylon Mixed Esters". The molecular weight of this sample is 23.55 Mw/10**5. The main chain is b13. Side chain 1 is Ac . Side chain 2 is La . The C number, branch and Degree of Substitution for acyl1 are: 2.0, 0.0 and 1.59 . The C number, branch and Degree of Substitution for acyl2 are: 12.0, 0.0 and 1.0 . The total DS is 2.59. Mw/Mn : 1.28, Tg(DSC)2 : 74.1, film/fiber : film, film/fiber type : hot-press, temperature for melting : 200.0, Tensil Strength(Mpa) : 6.5 +/- 19.9, Elongation at Break(%) : 6.9 +/- 3.2, Young's modulus(Gpa) : 0.6551 +/- 0.2389, Young's modulus(Gpa).1 : 0.434 +/- 0.0343, main : b13, link
AQA: 1) 0.8 2) 0.3 3) 1.0 4) 23.55 5) 0.5 6) 0.2 7) 0.0
205_PaAcPr
219_PaMy
215_PaMy
201_PaA
149_a13tBuAc
304_CDiBU
31_CdMy
2_CiHe
305_CDPi
308_CDiEn
307_CDiHe
302_CDHe
226_CdMy
230_CdMy
220_CdLa
301_CDVa
218_PaMy
ZoomPanSelect
Lasso
Zoom Fit
0
5
10
15
20
25
30
35
40
45
50
What is the molecular weight of this polymer sample in Mw/10**5?

Compositional Generation with Diffusion Probabilistic Models

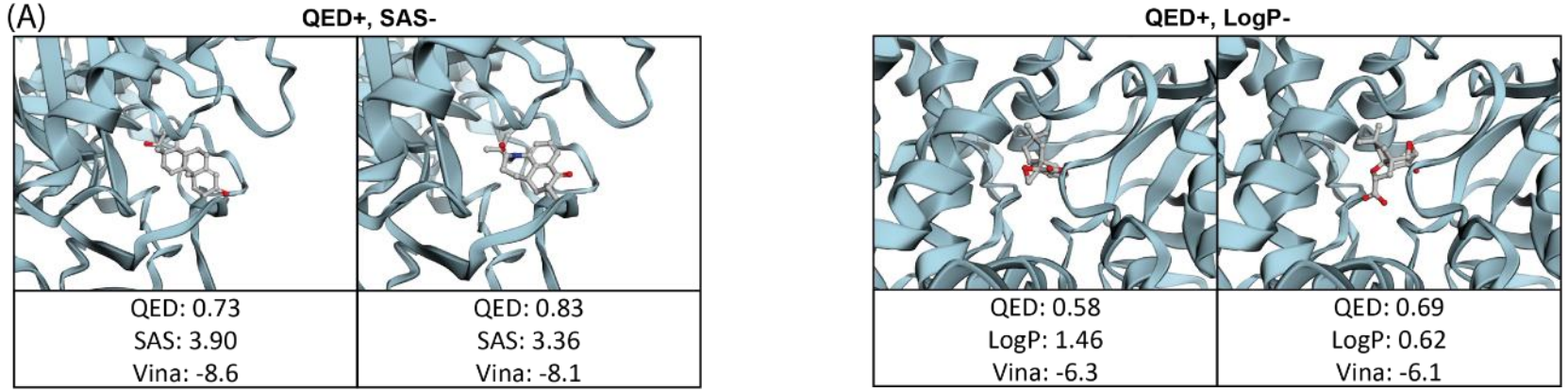
Compositionally manipulate multiple attributes in generation



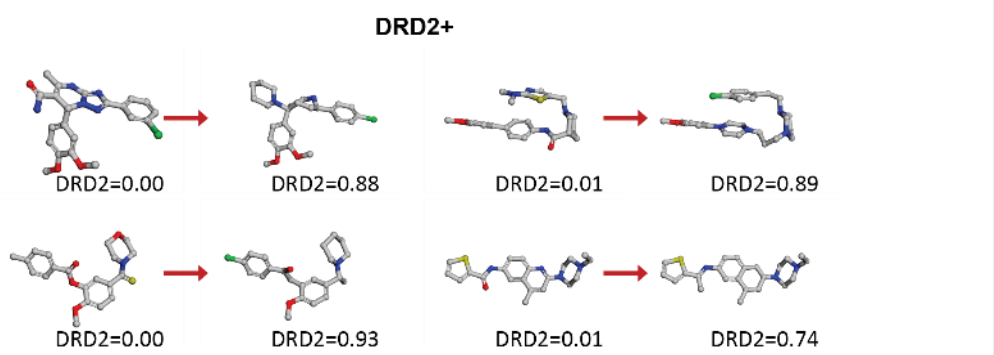
Shi *et al.*, Exploring Compositional Visual Generation with Latent Classifier Guidance. CVPR Workshop 2023.

Efficient Multi-Objective Manipulation of 3D Molecules

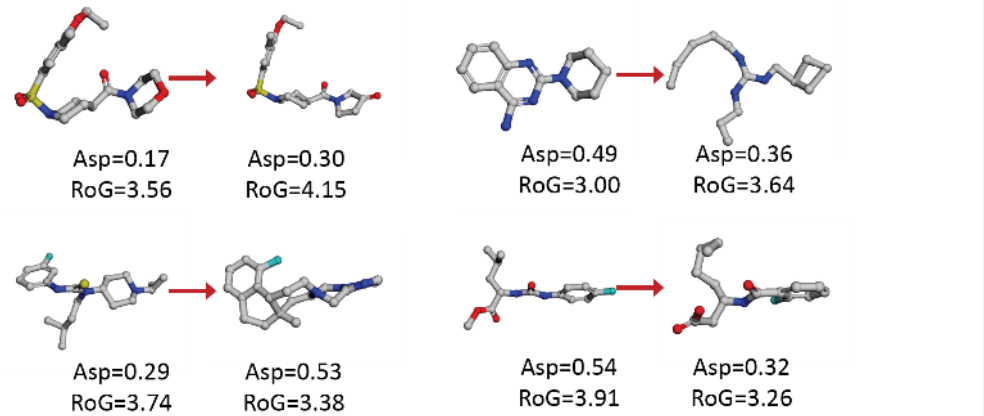
(1) Binding-preserving manipulation



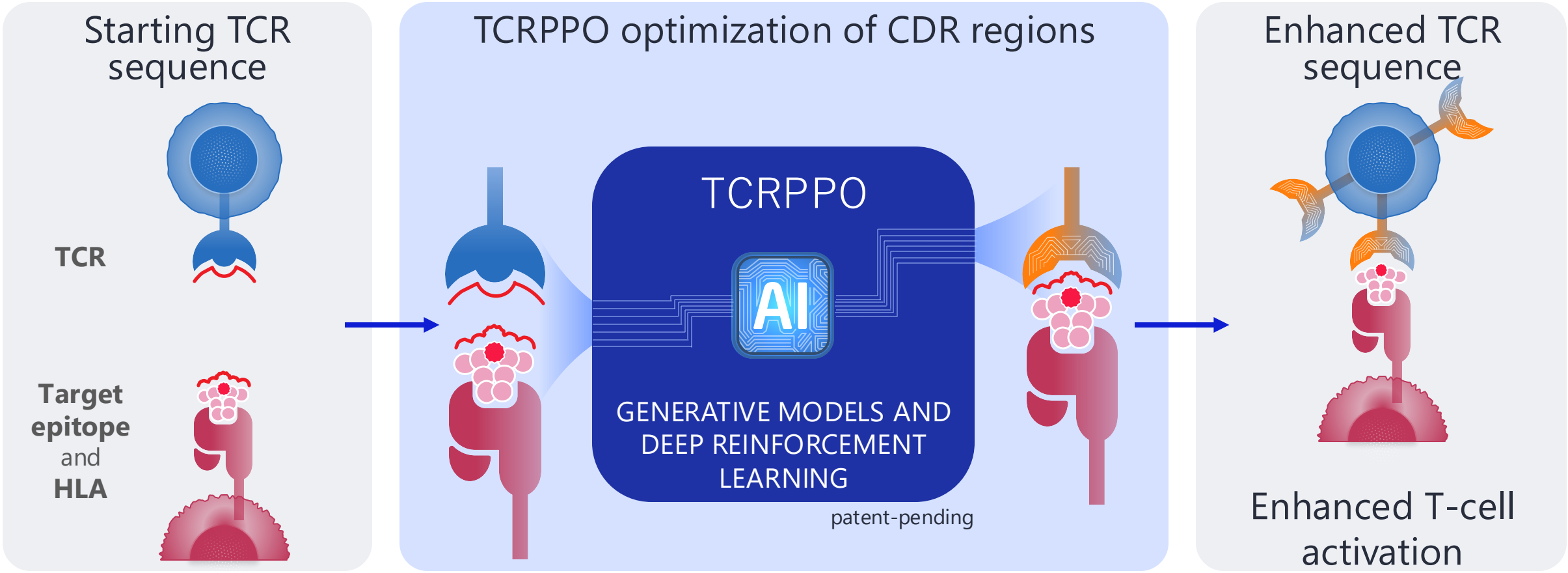
(2) Data efficient binding affinity enhancement (~1000 positive training data)



(3) Manipulation of 3D shapes and properties



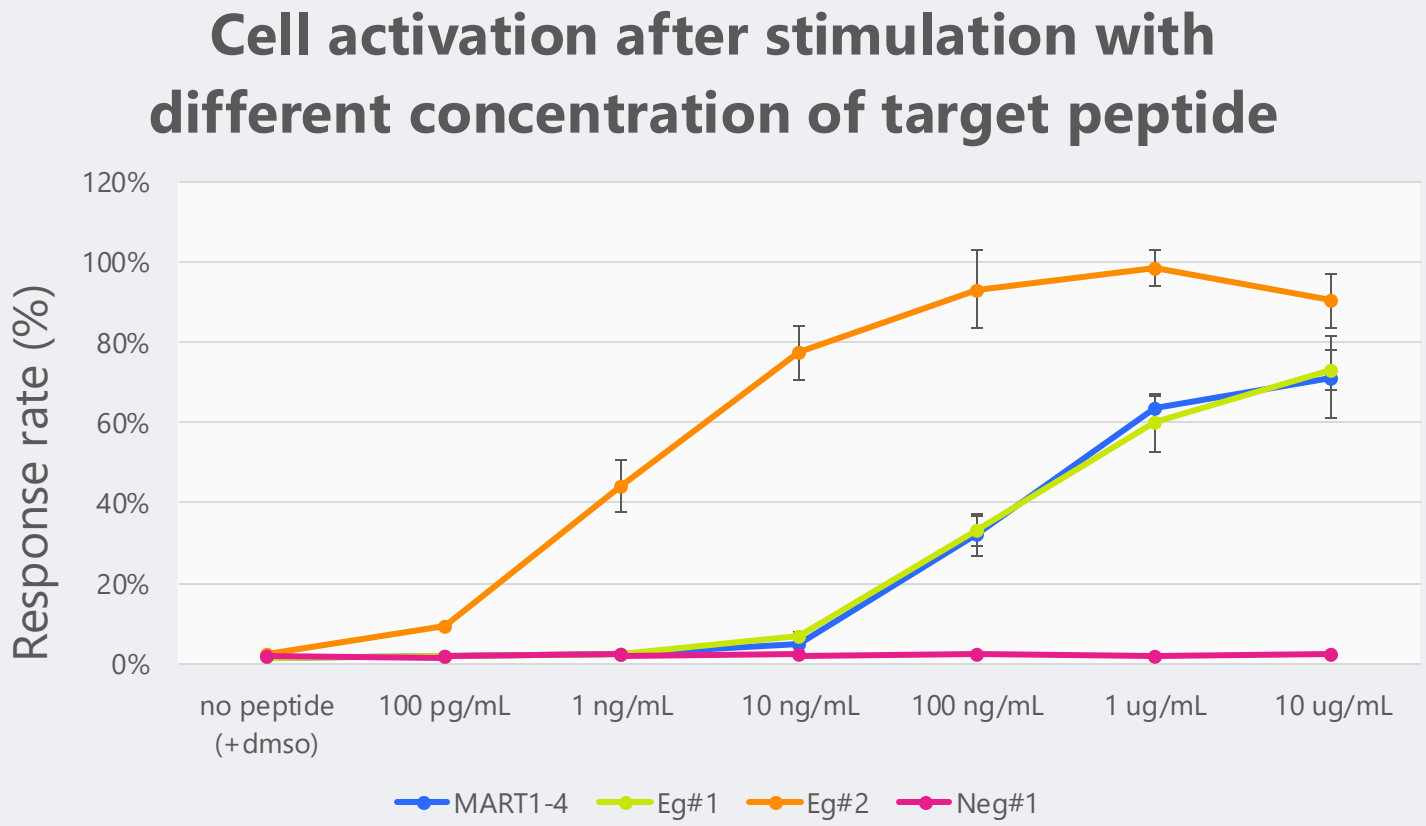
Generative Models and RL for TCR Engineering



Chen *et al.*, T-Cell Receptor Optimization with Reinforcement Learning and Mutation Polices for Precision Immunotherapy. RECOMB 2023.
Min *et al.*, Design of enhanced TCR against cancer antigens using an AI system. SITC 2024.

In Vitro Validation Results

Mutate existing TCRs for better immunotherapy with reinforcement learning



From over 6.4 billion mutated TCR candidates, we selected 2 optimized TCRs targeting the melanoma antigen MART-1 for wet-lab validation, one engineered TCR is over 1000 times more sensitive than the given template.

PPDiff: Diffusing in Hybrid Sequence-Structure Space for Protein-Protein Complex Design

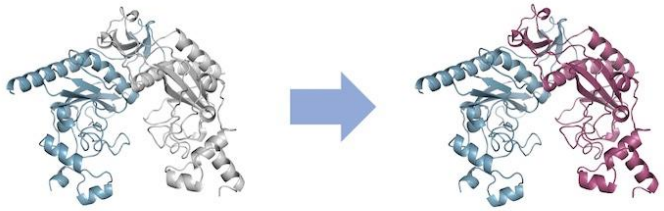
Language Technologies Institute | Carnegie Mellon University School of Computer Science

Zhenqiao Song, Tianxiao Li, Lei Li, Martin Renqiang Min

NEC Laboratories America

Code Link

Protein-Protein Complex Design

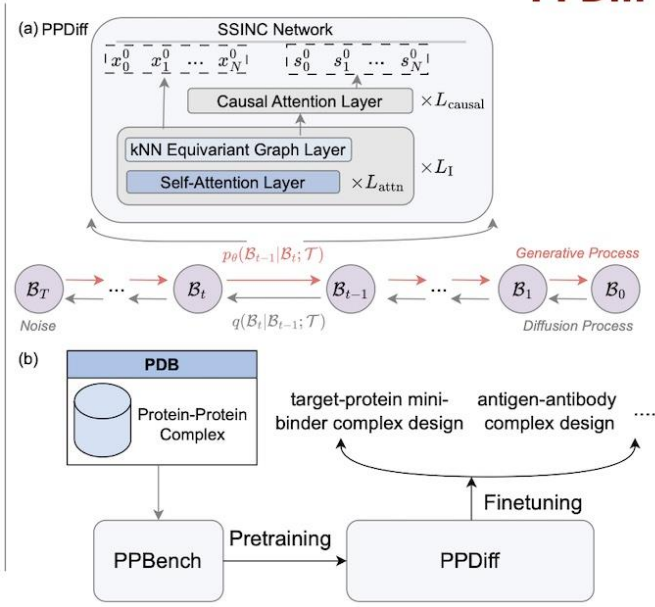


Given a target protein with its amino acid sequence and the alpha-carbon backbone structure, the goal is to designing proteins that bind to the target

Core idea

- Conditioning on target protein
- Co-designing binder protein sequence and backbone structure
- Hybrid (discrete-continuous) diffusion

PPDiff



B denotes binder protein

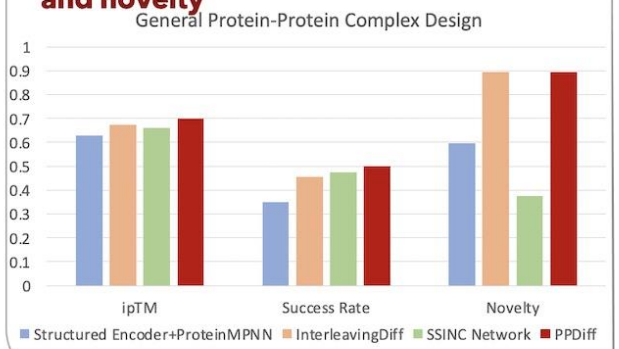
- x: alpha-carbon based backbone structure
- s: amino acid sequence

L_{attn}: 11 L_l: 3 L_{causal}: 1

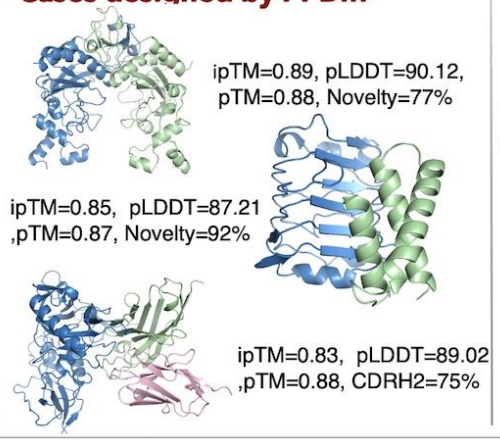
- PPBench
  - 734,032 complexes
    - Training: 706149, Validation: 155, Test: 56
- Target Protein-Mini Binder Complex
  - 10 target proteins
    - 5 seen classes, 5 zero-shot
- Antigen-Antibody Complex
  - 4261 complexes

General Protein-Protein Complex Design

PPDiff achieves the highest success rate and novelty

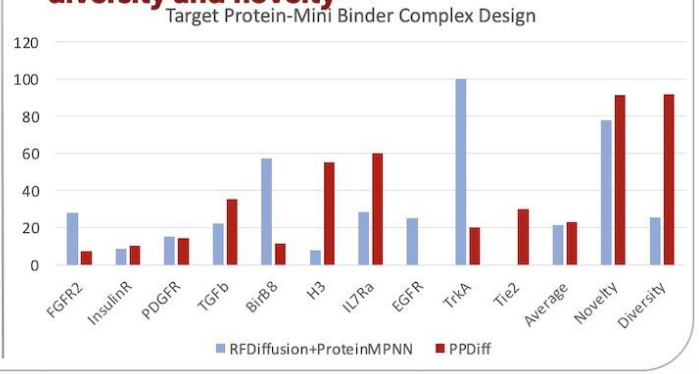


Cases designed by PPDiff



Target Protein-Mini Binder Complex Design

PPDiff achieves the higher success rate, diversity and novelty

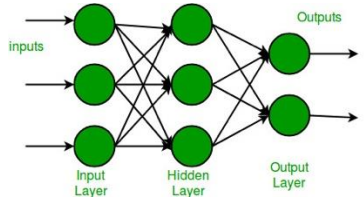


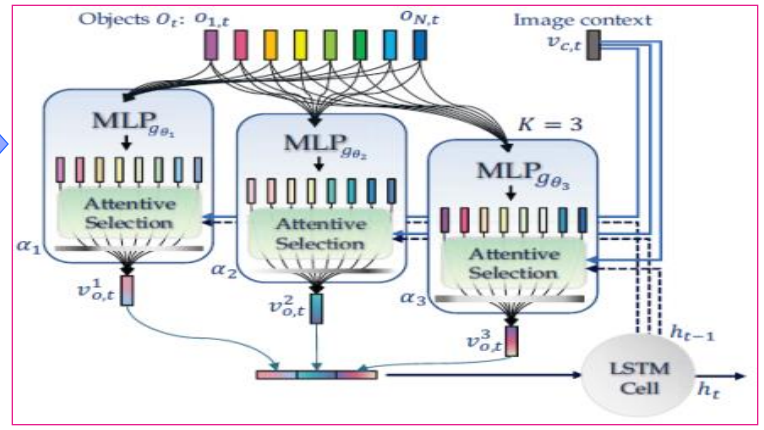
Future Directions

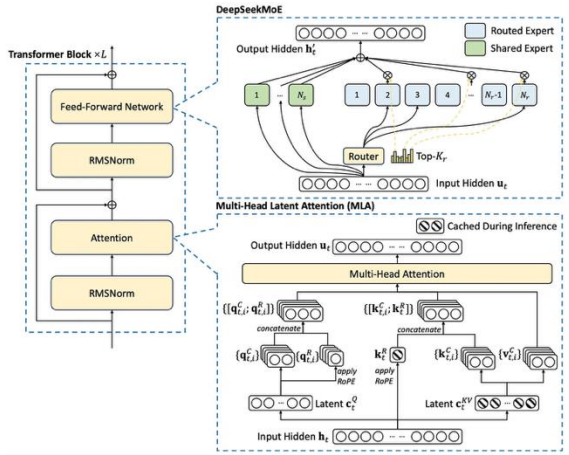
AI will make precision medicine scalable, accessible, and affordable worldwide.

What should we do?

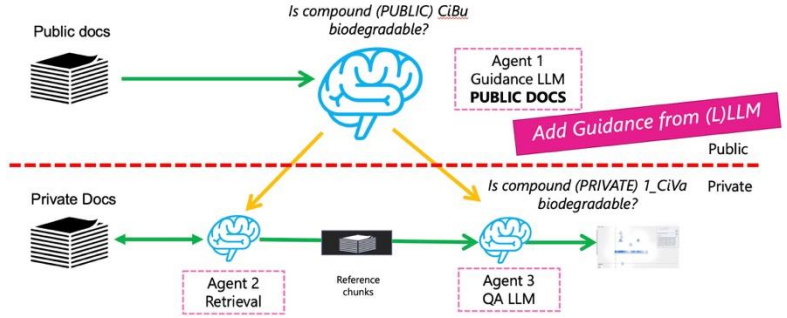
Capture Tacit Knowledge into Workflows: From LLM to Structured Workflow AI

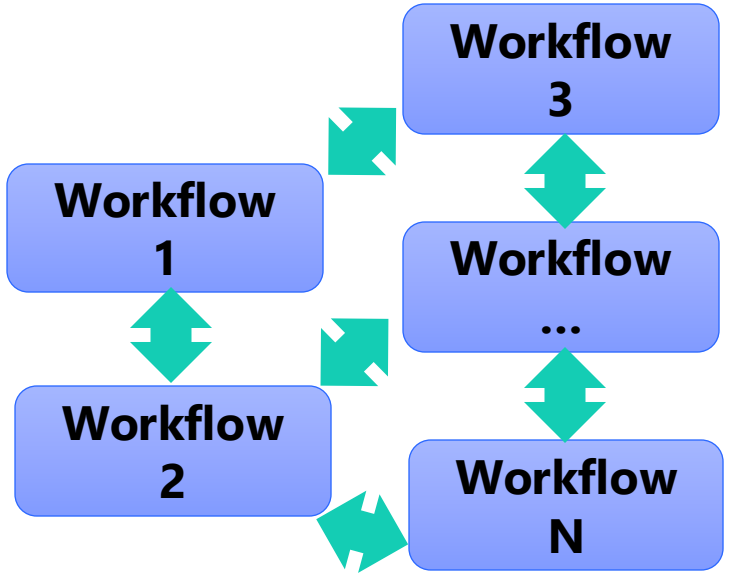






LLM





Digital Cell and Patient

Increasing efforts from both **pharmaceutical and IT companies** in building and utilizing digital twins of cells and patient models to reduce the trial-and-error of clinical research

sanofi

Homepage > Stories > Our Science > Digital "Twinning": Clinical Trials Powered by AI

Digital "Twinning": Clinical Trials Powered by AI

Published on: May 22, 2024

Twin Health multiplies funding with $50M for metabolic disease-reversing tech

By **Andrea Park** · Dec 14, 2023 12:45pm

Funding Summary
Last Round of Funding:
Series D **$50M** in Dec 2023
Lead Investor: Temasek Holdings
Total Funding: **$248.5M**

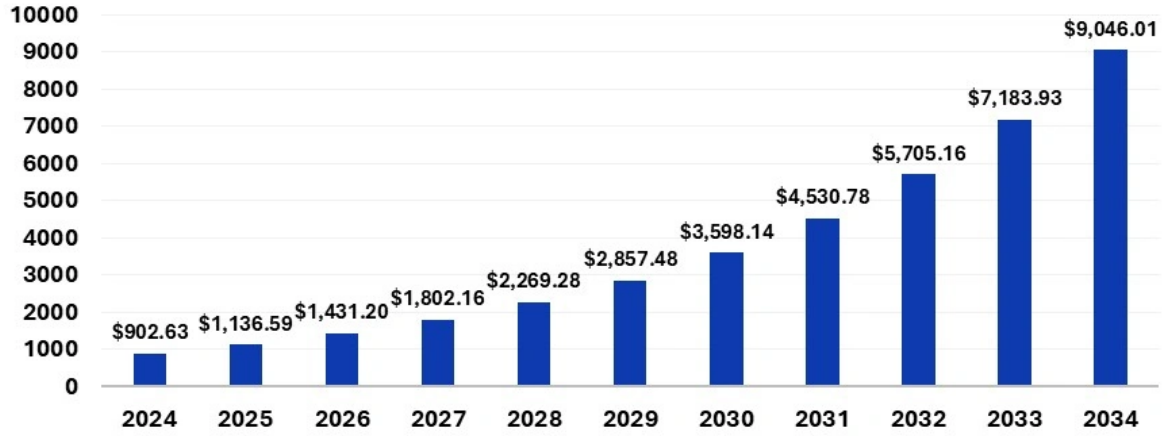
Digital twin generator Unlearn nets $50M to bolster clinical trials with AI models

By **Conor Hale** · Apr 19, 2022 12:24pm

Funding Summary
Last Round of Funding:
Series C **$50M** in Feb 2024
Lead Investor: Altimeter Capital
Total Funding: **$134.9M**

Precedence RESEARCH

Healthcare Digital Twins Market Size 2024 to 2034 (USD Million)



The global healthcare digital twins market size is predicted to increase from USD 902.63 million in 2024 to approximately USD 9,046.01 million by 2034, expanding at a CAGR of 25.92% from 2025 to 2034.

Source: https://www.precedenceresearch.com/healthcare-digital-twins-market

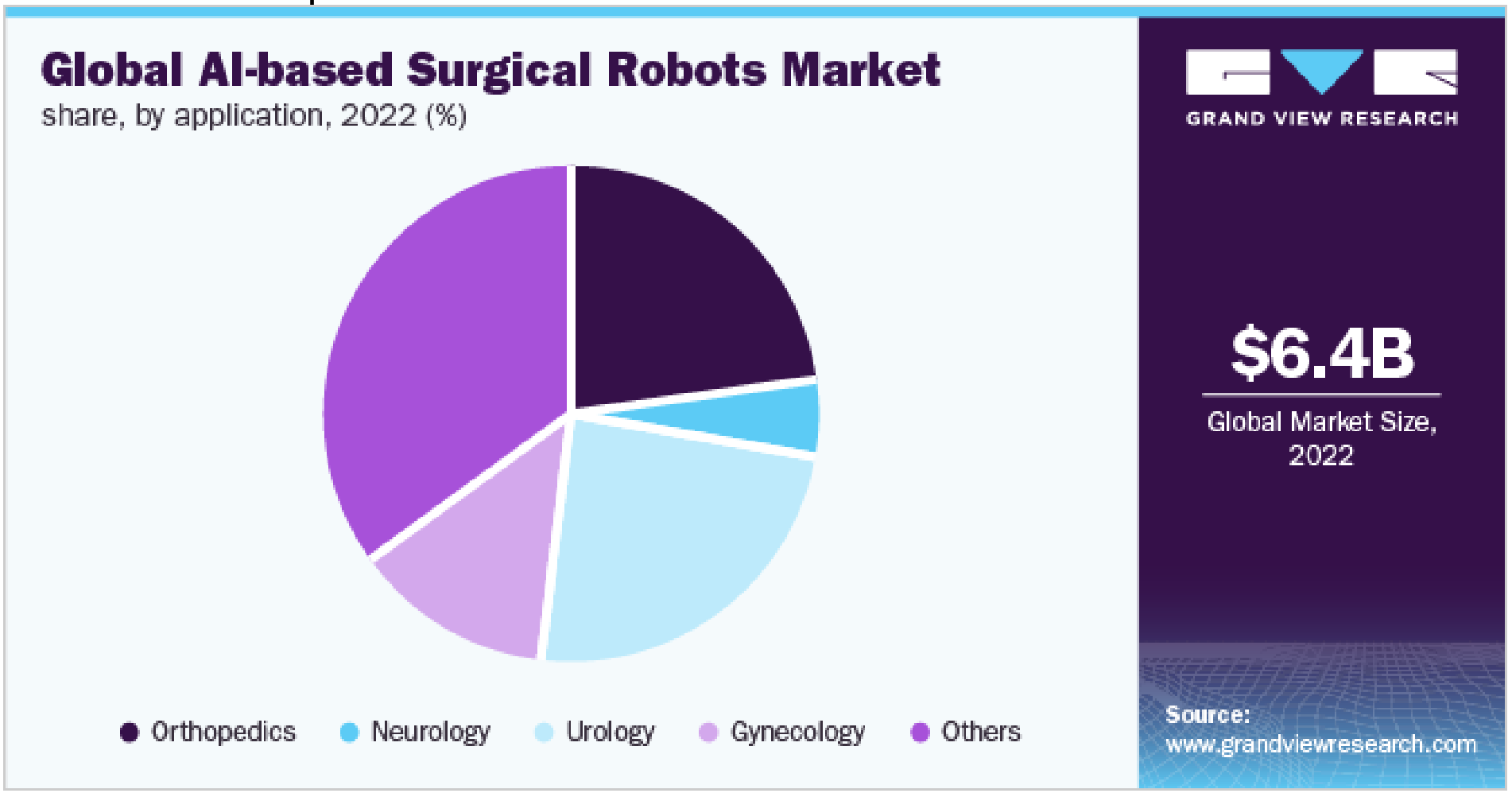
Connected Health

Primary care providers, virtual care providers, and **specialist clinics** will buy services to summarize incoming medical records, triage inquiries from patients and augment doctor-patient interaction

https://www.grandviewresearch.com/industry-analysis/connected-health-wellness-devices-market

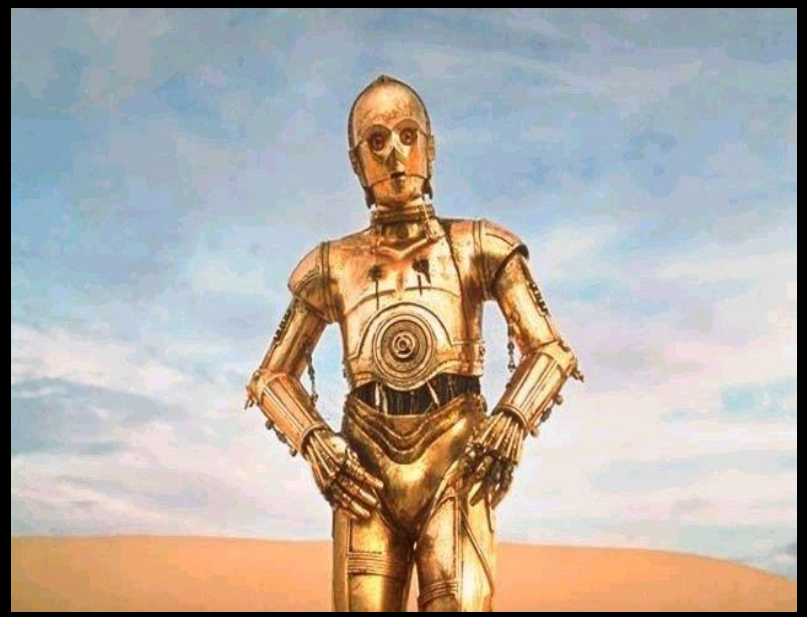
AI-driven Robotic Surgery

AI-based surgical robots' segment was valued at $6.4 billion in 2022 and is expected to expand at a CAGR of **18.9% from 2023 to 2030**



The Ultimate Future of Machine Learning with "Simulated Consciousness"

From Generative Digital Cell and Digital Human to Mike Peter Robot



Bill Gates Professor of Medicine, Molecular Biophysics & Biochemistry and Computer Science
A.B. 2039, Harvard University
Ph.D. 2043, Cambridge University
Joined Yale Faculty 2047

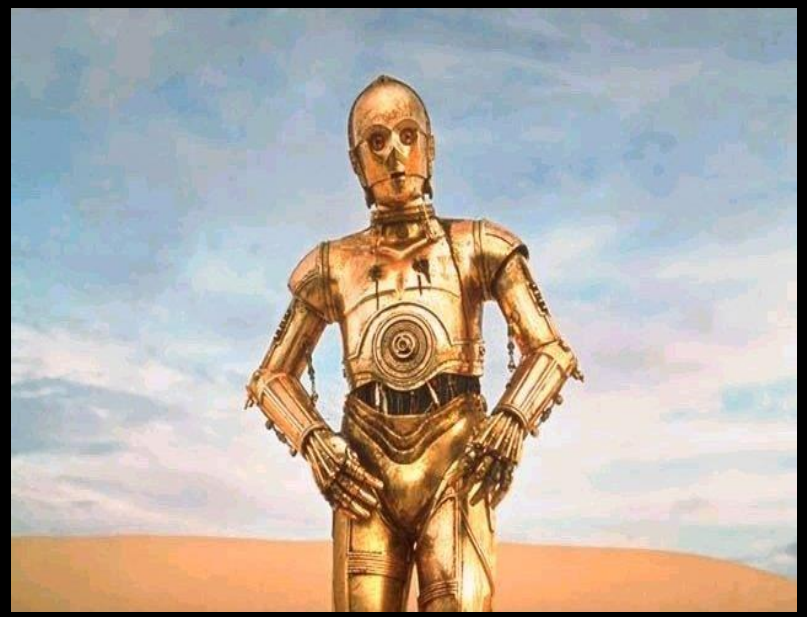
Professor Robot does research in the new field of Rapid Human Disease Treatment and Prevention. He won the Nobel Prize in Physiology for discovering ABCBBCNBC007 for the no-pain one-smell treatment of cancer in 2067.

Most of this slide was shown during a talk at Yale in 2015

The Ultimate Future of Machine Learning with "Simulated Consciousness"

From Generative Digital Cell and Digital Human to

Of the Human, By the Human, For the Human



Mike Peter Robot

Bill Gates Professor of Medicine, Molecular Biophysics & Biochemistry and Computer Science
A.B. 2039, Harvard University
Ph.D. 2043, Cambridge University
Joined Yale Faculty 2047

Professor Robot does research in the new field of Rapid Human Disease Treatment and Prevention. He won the Nobel Prize in Physiology for discovering ABCBBCNBC007 for the no-pain one-smell treatment of cancer in 2067.

Most of this slide was shown during a talk at Yale in 2015

Acknowledgements

NEC Labs and NEC Bio

Kaidre Bendjama
Trevor Clancy
Pait Ivy Grace
Tianxiao Li
Pierre Machart
Christopher Malon
Iain Melvin
Andrea Meiser
Daiki Mori
Anja Moesch
Kaz Onoguchi
Jonathan Warrell

Interns and Students

Ziqi Chen (OSU)
Haoran Liu (TAMU)
Junhao Liu (UCI)
Youzhi Luo (TAMU)
Haomiao Ni (PSU)
Yiren Jian (Dartmouth)
Xuyang Shen (UConn)
Changhao Shi (UCSD)
Zhenqiao Song (CMU)

Collaborators

Prof. Chao Cheng (Baylor)
Prof. Mark Gerstein (Yale)
Prof. Hongyu Guo (Ottawa)
Prof. Lei Li (CMU)
Prof. Xia Ning (OSU)
Prof. Dongjin Song (UConn)
Prof. Jing Zhang (UCI)

\Orchestrating a brighter world