Unleashing Creativity Using AI Agent Networks

Mihaela van der Schaar, FRS

John Humphrey Plummer Professor of Machine Learning, Artificial Intelligence and Medicine, University of Cambridge
Director, Cambridge Center for AI in Medicine

van_der_Schaar \ LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

mv472@cam.ac.uk

@MihaelaVDS

linkedin.com/in/mihaela-van-der-schaar/

Meet our group!

https://www.vanderschaar-lab.com/
→ Research Team

■ = joined us in 2025


Anita Kriz


Antonin Berthon

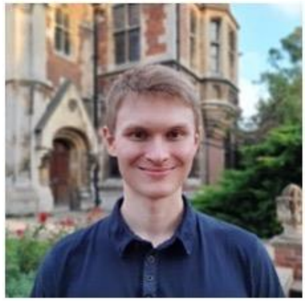

Benjamin Lapostolle


Claudio Fanconi


Harry Amad

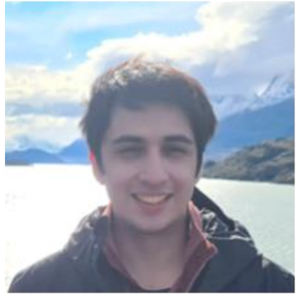

Julianna Piskorz

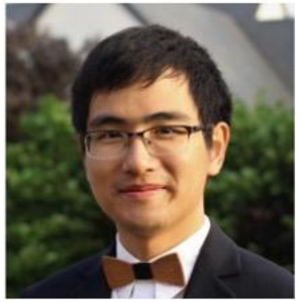

Kasia Kobalczyk


Krzysztof Kacprzyk


Luca Muscarnera


Max Ruiz Luyten


Nicolás Astorga


Nicholas Huyn


Paulius Rauba


Qiyao Wei

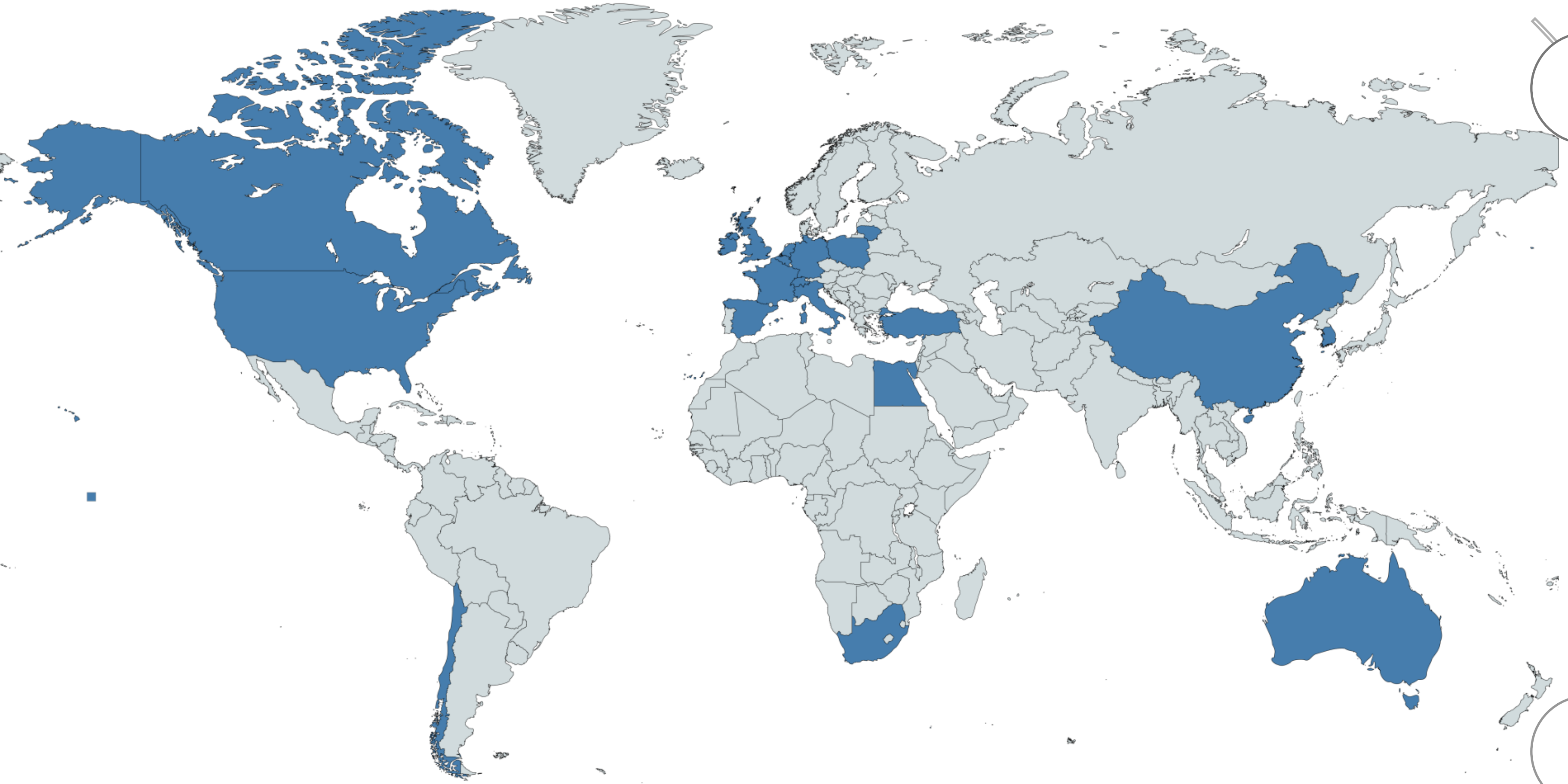

Tennison Liu


Thomas Pouplin


Victor Baillet

Our lab – diverse and international

From over 20+ countries, from diverse academic backgrounds



- Economists
- Statisticians
- Mathematicians
- Computer Scientists
- Physicists
- Engineers

A fundamentally new paradigm is needed for AI to orient it towards addressing the complexities of the real-world

What is Reality-centric AI?

- AI which aims to solve real-world problems
- AI which operates effectively and accountably given the inherent and unavoidable complexities of the real world
- AI which empowers, and does not marginalize humans

Our focus

Identify big and bold problems we want to solve – societal, healthcare, education, finance, economics, biology etc.

Define and formalize the identified problems

Invent cutting-edge ML solutions to solve them

But there is a gap….because most of these problems are complex & not tractable

This also requires out-of-the-box thinking, modeling and new ML to close the gap

Fearless thinking

We are interested to envision a better world and then develop new ways of thinking and new ML to conceive it

A huge challenge!
Brain-hurting difficult!
Exciting!
Intellectual adrenaline ☺
A sense of achievement

Different types of research labs

Most ML/AI Research Labs

- Work on well-formalized problems, well-known benchmarks
- Focus on the same problem for many years
- Build on established work
- Use similar methods repeatedly
- All work in the same sub-area of ML
- Create excitement through the development of new algorithmic possibilities

Our way

- Focus on transformation in the real-world, create new ways of thinking, newly defined problems
- Formalize new, unique challenges
- Different problems require diverse methods and new strands of ML
- Explore multiple areas of ML, beyond a single sub-area
- Create excitement through groundbreaking possibilities

ML topics

AI Agents: Creativity, Reasoning, Safety, Collaboration, Labor Markets
AI for Scientific Discovery
Synthetic data, simulators and generative models
Digital Twins
Decision making under uncertainty
Next-generation RL
Causal Inference
Time series
AI-human alignment, AI for human empowerment
AI for complex systems, organizations
Autoformalism
AI for Operations Research
AI for Education
Etc.

van der Schaar lab @ NeurIPS, ICML, ICLR 2025

https://www.vanderschaar-lab.com/
→ Publications

32 papers published in 2025 at the best AI/ML conferences

ICLR

van_der_Schaar
\LAB

van der Schaar lab @ ICLR 2025

Risk-Sensitive Diffusion: Robustly Optimizing Diffusion Models with Noisy Samples
Y. Li, M. R. Luyten, M. van der Schaar

Going Beyond Static: Understanding Shifts
... Attribution
... M. van der Schaar

...uction via Semantical
...n der Schaar

No Equations Needed: Learning System Dynamics Without Relying on Closed-Form ODEs
K. Kacprzyk, M. van der Schaar

Active Task Disambiguation with LLMs
K. Kobalczyk, N. Astorga, T. Liu, M. van der Schaar

ICML
International Conference
On Machine Learning

van_der_Schaar
\LAB

van der Schaar lab @ ICML 2025

Truly Self-Improving Agents Require Intrinsic Metacognitive Learning
T. Liu, M. van der Schaar

Not All Explanations for Deep Learning Phenomena Are Equally Valuable
A. Jeffares, M. van der Schaar

All Current Generative Fidelity and Diversity Metrics are Flawed
O. Räisä · B. van Breugel, M. van der Schaar

Statistical Hypothesis Testing for Auditing Robustness in Language Models
P. Rauba, Q. Wei, M. van der Schaar

Strategic Planning: A Top-Down Approach to Option Generation
M. R. Luyten, A. Berthon, M. van der Schaar

Bootstrapping Self-Improvement of Language Model Programs for Zero-Shot Schema Matching
N. Seedat, M. van der Schaar

Continuously Updating Digital Twins using Large Language Models
H. Amad, N. Astorga, M. van der Schaar

Unified Screening for Multiple Diseases
Y. Narter, A. Hüyük, M. van der Schaar, C. Tekin

Skip the Equations: Learning Behavior of Personalized Dynamical Systems Directly From Data
K. Kacprzyk, J. Piskorz, M. van der Schaar

Preference Learning for AI Alignment: a Causal Perspective
K. Kobalczyk, M. van der Schaar

AutoCATE: End-to-End, Automated Treatment Effect Estimation
T. Vanderschueren, T. Verdonck, M. van der Schaar, W. Verbeke

Stochastic Encodings for Active Feature Acquisition
A. Norcliffe, C. Lee, F. Imrie, M. van der Schaar, P. Lió

Autoformulation of Mathematical Optimization Models Using LLMs
N. Astorga, T. Liu, Y. Xiao, M van der Schaar

G-Sim: Generative Simulations with Large Language Models and Gradient-Free Calibration
S. Holt, M. R. Luyten, A. Berthon, M. van der Schaar

The Synergy of LLMs & RL Unlocks Offline Learning of Generalizable Language-Conditioned Policies with Low-fidelity Data
T Pouplin, K Kobalczyk, H Sun, M van der Schaar

NEURAL INFORMATION
PROCESSING SYSTEMS

van_der_Schaar
\LAB

van der Schaar lab @ NeurIPS 2025

Improving the Generation and Evaluation of Synthetic Data for Downstream Medical Causal Inference
Harry Amad*, Zhaozhi Qian, Dennis Frauen, Julianna Piskorz, Stefan Feuerriegel, Mihaela van der Schaar

Towards a Cascaded LLM Framework for Cost-effective Human-AI Decision-Making
Claudio Fanconi*, Mihaela van der Schaar

Treatment Effect Estimation for Optimal Decision-Making
Dennis Frauen*, Valentyn Melnychuk, Jonas Schweisthal, Mihaela van der Schaar, Stefan Feuerriegel

Timely Clinical Diagnosis through Active Test Selection
Silas Ruhrberg Estévez*, Nicolás Astorga, Mihaela van der Schaar

Semantic-KG: Using Knowledge Graphs to Construct Benchmarks for Measuring Semantic Similarity
Qiyao Wei*, Edward Morrell, Lea Goetz, Mihaela van der Schaar

Simulating Viva Voce Examinations to Evaluate Clinical Reasoning in Large Language Models
Christopher Chiu*, Silviu Pitis, Mihaela van der Schaar

PhD Opportunities at the van der Schaar Lab

We are looking for 4 fully funded PhDs to join our lab! Applications now OPEN!

About our PhD programme

- Based at the **University of Cambridge (DAMTP)**
- Work at the cutting edge in a world-leading lab
- Projects with a purpose; work that can change the world
- **32** papers accepted at four largest AI/ML conferences (NeurIPS, ICML, ICLR, AISTATS) in the past year

Apply through our website:
www.vanderschaar-lab.com/join-the-van-der-schaar-lab/

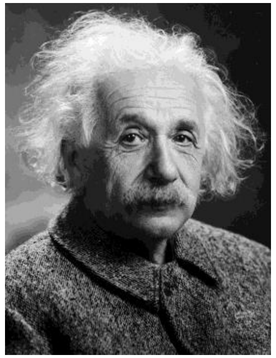
For any queries, please email:
vanderschaarlab@damtp.cam.ac.uk

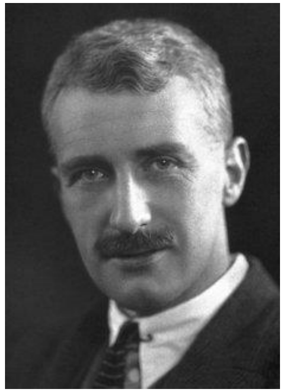
Join our PhD Open Day

- **Monday 20 October, 9:30am (BST)** *(virtual)*
- Hear from current PhD students, explore research areas, and ask questions

Scan the QR code to register for the PhD Open Day:

Scientific discovery in the era of AI





- If Einstein had access to powerful AI, would he still have needed intuitive leaps or would AI have led him directly to relativity?
- Could AI agents reveal laws of nature that even the greatest human minds (Einstein, Hill, Newton) could never have grasped alone?
- In an era of AI-driven discovery, do scientists remain the explorers or are we becoming interpreters of insights unearthed by our digital counterparts?
- Will creativity in science shift from deriving answers to framing the most meaningful questions?
- Are we approaching an age where discovery emerges from symbiotic human–AI partnerships where AI collaborates with us, rather than replace us?
- And if so, what does it mean to be a scientist when knowledge itself can now be generated, reasoned, and evolved by AI agents?

This talk

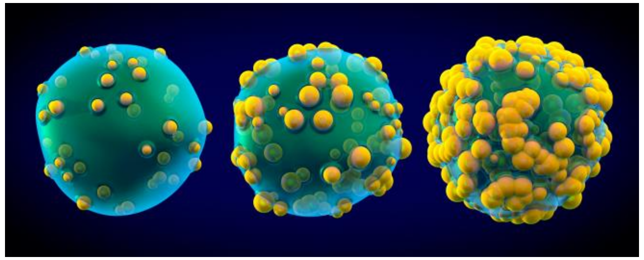
We are entering a new **paradigm of discovery,** driven not by static models or datasets, but by **AI agents** capable of reasoning, experimentation, and evolving knowledge, transforming science from human-centric to **co-creative**

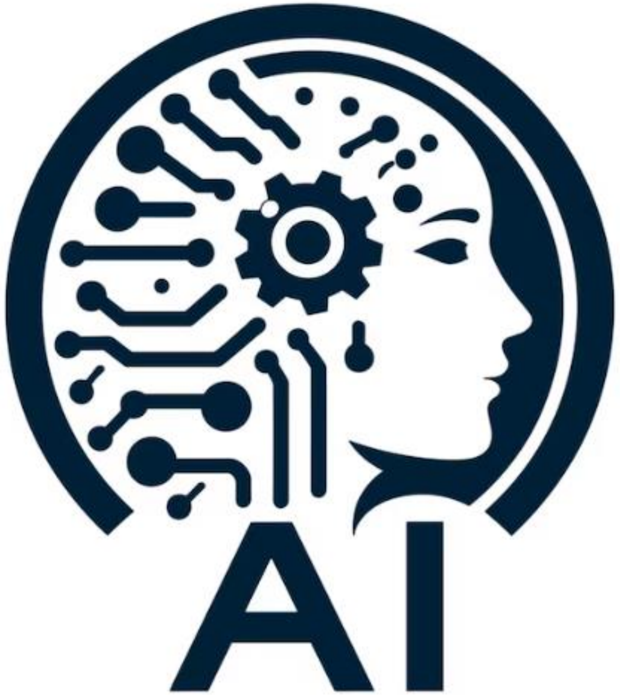
- LLM-based AI Agents
  - Why are they exciting
  - Examples
- Genies: Next-generation AI Agents
- Networks of Autonomous Scientific Agents
- Join forces with us

Discovering scientific laws

- Uncovering the Fundamental Laws Governing System Evolution Over Time
- Why do we care?
- Applications: Medicine, Physics, Finance, Climate, etc.







Discovery of governing equations using ML

	Explicit function	Ordinary differential equation	Partial differential equation
Typical form		$\frac{dx}{dt} = f(x, t)$	
Examples			

<u>A hard problem</u>

$$x(t) \xleftrightarrow{\text{ODE}} x'(t)$$

Why do we care about ODEs?

Almost everything in medicine is a dynamical system

1. **Pharmacokinetic (PK) Models**
2. **Pharmacodynamic (PD) Models**
3. **Physiological-Based Pharmacokinetic (PBPK) Models**
4. **Homeostatic Regulation Models and Non-Homeostatic models**
5. **Epidemic and Population Dynamics**
6. **Biochemical Pathway Models**
7. **Physiological Models**

To describe dynamical systems, we need

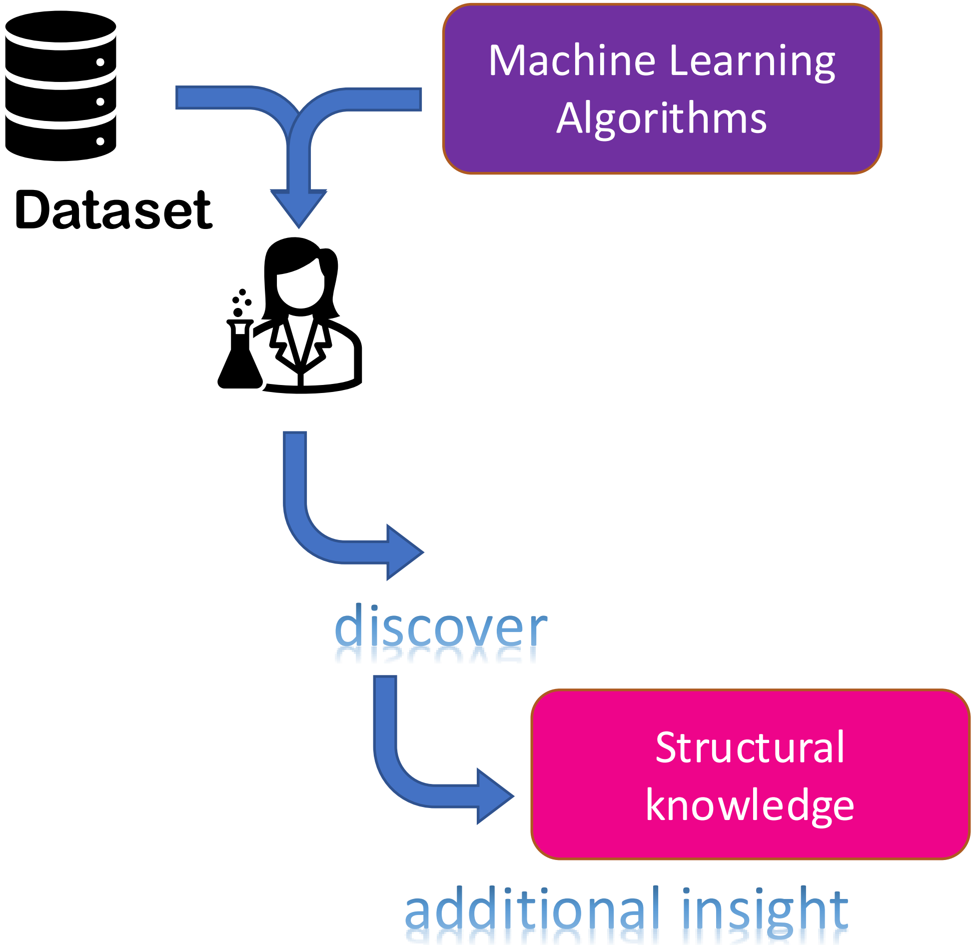
Differential equations

- Equations that involve derivatives
- Commonly used to describe continuous-time dynamical systems
- Describe the change in infinitesimal time (time derivative)
- E.g. Ordinary DE

$$x(t) \xleftrightarrow{\text{ODE}} x'(t)$$

Learning ODEs from data: A hard problem

Problem formulation



Dataset

$$\left\{ \begin{array}{c} \mathbf{y}_1(t) \\ \vdots \\ \mathbf{y}_D(t) \end{array} \right\}$$

$t \in \{t_1, t_2, \dots, T\}$

$\mathbf{y}_i(t) \in \mathbb{R}^J$

Goal: Discover

System of J ODEs

$$\left\{ \begin{array}{c} f_1(\boldsymbol{x}) = \dot{x}_1 \\ \vdots \\ f_J(\boldsymbol{x}) = \dot{x}_J \end{array} \right\}$$

$x_j : [0, T] \to \mathbb{R}$

$\boldsymbol{x}(t) = [x_1, \dots, x_J]^T$

$f_j : \mathbb{R}^J \to \mathbb{R}$

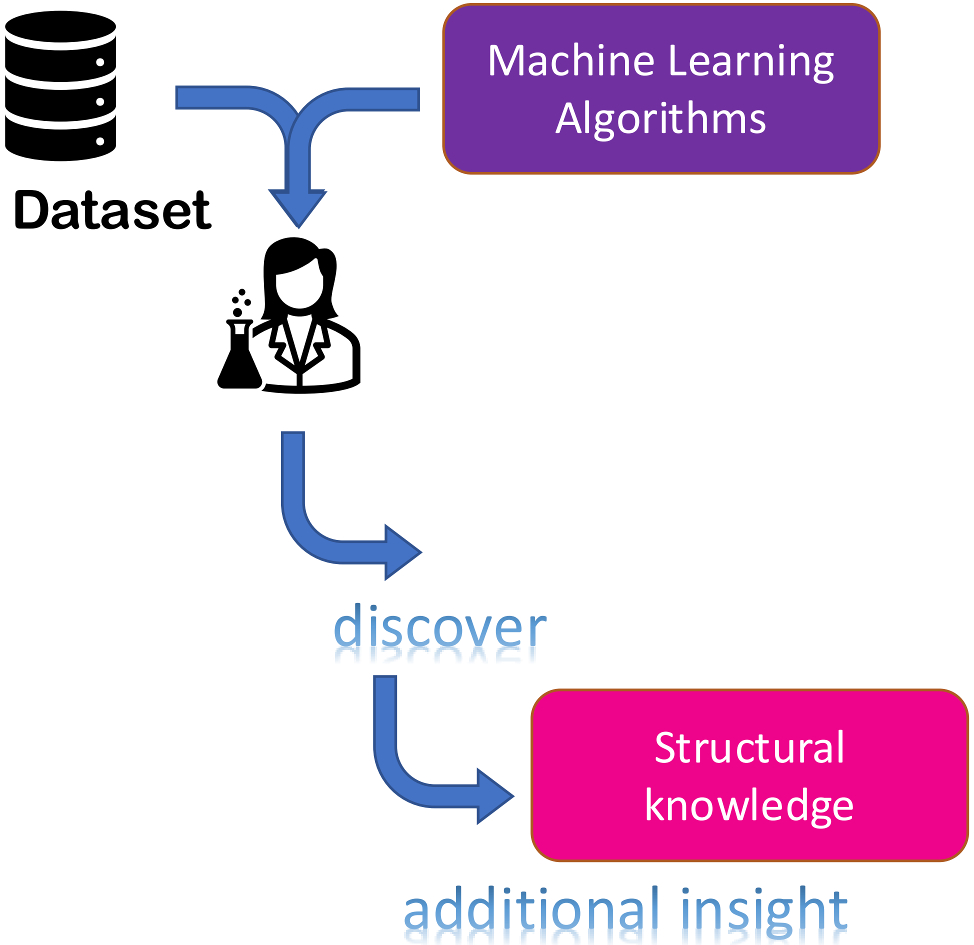
Why is discovering ODEs from data challenging?

1. The time derivative is not observed
    - Only observe the states over time
    - Conventional *symbolic regression* methods are not applicable

2. It is difficult to estimate the time derivative
    - States are observed sporadically with noise
    - Naïve two-step symbolic regression is likely to fail

3. Difficulty in directly solving the initial value problem of ODE
    - The true initial condition is unknown & difficult to infer
    - Sensitive to initial condition
    - Computationally challenging

Discover closed-form ordinary differential equations (ODEs) from observed trajectories - *D-CODE*



Z. Qian, K. Kacprzyk, M. van der Schaar, ICLR 2022

Zhaozhi Qian

Krzysztof Kacprzyk

D-CODE: theory

Variational formulation of ordinary differential equations

$$\dot{x}_j(t) = f_j(\boldsymbol{x}(t)),\ \forall j = 1, \dots, J,\ \forall t \in [0, T] \quad (1)$$

Definition 1. Consider $J \in \mathbb{N}^+$, $T \in \mathbb{R}^+$, continuous functions $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$, $f : \mathbb{R}^J \to \mathbb{R}$, and $g \in \mathcal{C}^1[0, T]$, where $\mathcal{C}^1$ is the set o … e define the functionals

$$C_j(f, \boldsymbol{x}, g) := \int_0^T f(\dots 1, 2, \dots, J\} \quad (3)$$

Theorem 1. *Consider $J \in \mathbb{N}^+$, $j \in \{$ … tinuous function, and let $\boldsymbol{x} : [0, T] \to \mathbb{R}^J$ be a conti … $(t) = f^*(\boldsymbol{x}(t))$. Consider a sequence of functions $(\widehat{\boldsymbol{x}}$ … ly differentiable function. If $(\widehat{\boldsymbol{x}}_k)$ converges to $\boldsymbol{x}$ in L … ction f*

$$\lim_{S\to\infty} \lim_{k\to\infty} \sum_{s=1} C_j(f, \widehat{\boldsymbol{x}}_k, g_s)^2 = d_{\boldsymbol{x}}(f, f^*)^2, \quad (7)$$

where $\{g_1, g_2, \dots\}$ is a Hilbert (orthonormal) basis for $L^2[0, T]$ such that $\forall i$, $g_i(0) = g_i(T) = 0$ and $g_i \in \mathcal{C}^1[0, T]$.

Natural choice

$$g_s(t) = \sqrt{2/T}\sin(s\pi t/T)$$

LLM Agents

LLM Agents: Beyond Text Generation - LLMs That Act

LLMs: Large Language Models that predict text based on input prompts.
LLM Agents: Systems that augment LLMs with:

- **Context management (memory)**
- **Decision-making (planning, goal-setting)**
- **Actions (tool usage, APIs, environment interaction)**

The LLM Agent Loop

Prompt & Context: The agent queries the LLM with a prompt that includes goals, instructions, data etc.
Reasoning: The agent reasons (e.g. CoT) considering multiple steps or sub-tasks.
Action: The agent can execute actions (e.g., call a tool, fetch external info, revise the approach).
Feedback: The agent updates its memory/context with the results of the action.
Iteration: Cycle repeats until the goal is reached (or stopping criterion is met).

Take away: LLM is no longer just a passive language predictor; it's part of a loop that actively pursues objectives.

D3: A Discovery Agent for Transparent Dynamical Laws

Data-Driven Discovery of Dynamical Systems using Large Language Models

NeurIPS 2024, Spotlight


Samuel Holt

Zhaozhi Qian


Tennison Liu


Jim Weatherall


Mihaela van der Schaar

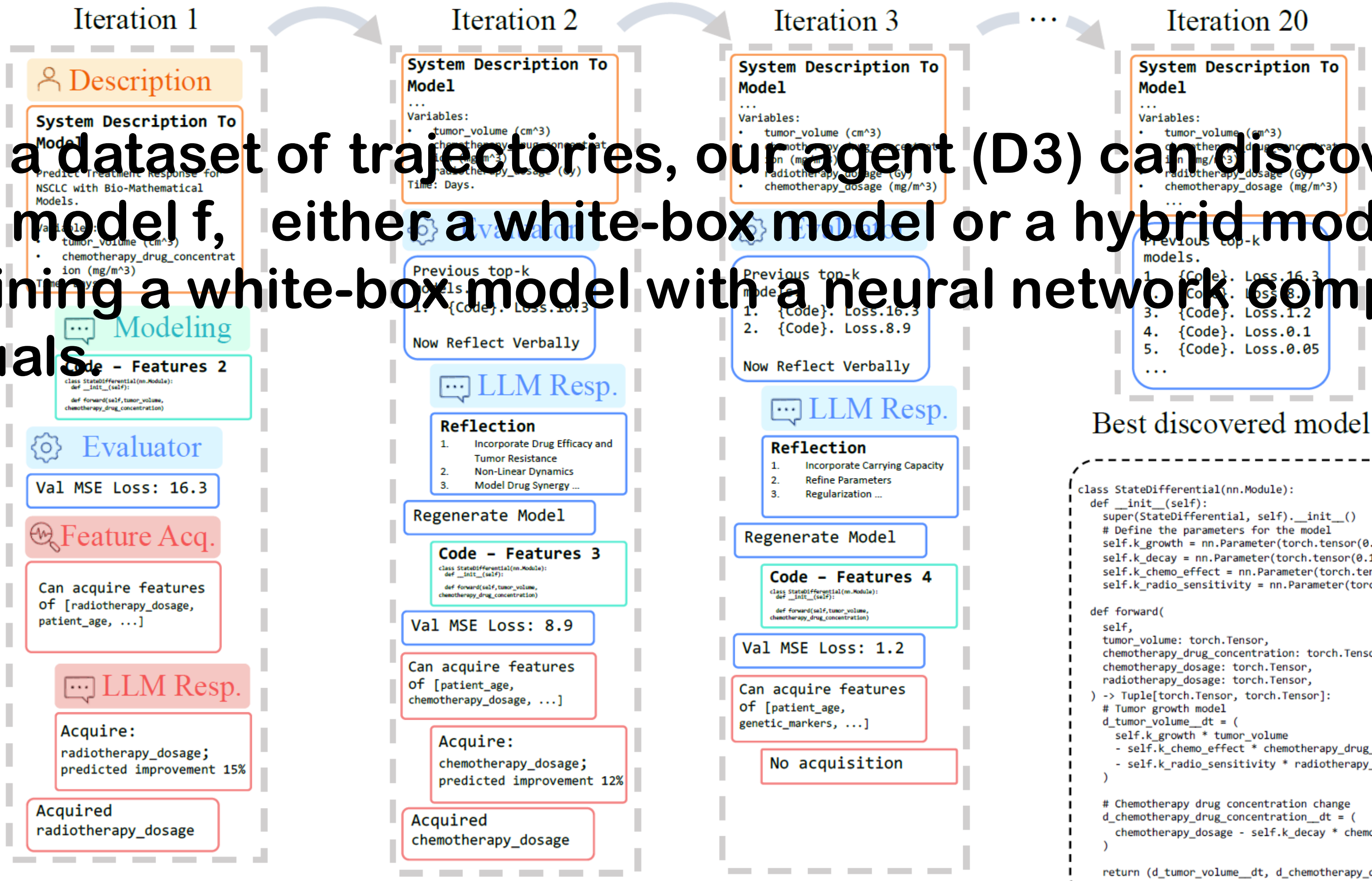
van_der_Schaar \LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

Our agent iteratively discovers and refines system dynamics

Given a dataset of trajectories, our agent (D3) can discover a well-fitting model f, either a white-box model or a hybrid model, combining a white-box model with a neural network component fit to residuals.




```
class StateDifferential(nn.Module):
  def __init__(self):
    super(StateDifferential, self).__init__()
    # Define the parameters for the model
    self.k_growth = nn.Parameter(torch.tensor(0.1))
    self.k_decay = nn.Parameter(torch.tensor(0.1))
    self.k_chemo_effect = nn.Parameter(torch.tensor(-0.1))
    self.k_radio_sensitivity = nn.Parameter(torch.tensor(-0.1))

  def forward(
    self,
    tumor_volume: torch.Tensor,
    chemotherapy_drug_concentration: torch.Tensor,
    chemotherapy_dosage: torch.Tensor,
    radiotherapy_dosage: torch.Tensor,
  ) -> Tuple[torch.Tensor, torch.Tensor]:
    # Tumor growth model
    d_tumor_volume__dt = (
      self.k_growth * tumor_volume
      - self.k_chemo_effect * chemotherapy_drug_concentration * tumor_volume
      - self.k_radio_sensitivity * radiotherapy_dosage * tumor_volume
    )

    # Chemotherapy drug concentration change
    d_chemotherapy_drug_concentration__dt = (
      chemotherapy_dosage - self.k_decay * chemotherapy_drug_concentration
    )

    return (d_tumor_volume__dt, d_chemotherapy_drug_concentration__dt)
```

New Discovered PK Warfarin Model

Experiments on a real pharmacokinetic Warfarin dataset

- D3 uncovers a new plausible pharmacokinetic model
- Outperforms existing literature
- Highlighting its potential for precision dosing in clinical applications

$$\frac{dC}{dt} = \sqrt{D} - k_{\text{eff}} \cdot \frac{C}{K_m + C},$$

$$\begin{aligned} k_{\text{eff}} = {} & k_{e,\text{base}} + k_{e,\text{age}} \cdot (A - \overline{A}) + k_{e,\text{sex}} \cdot (S - \overline{S}) \\ & + k_{\text{decay}} \cdot C + k_{ds} \cdot D \cdot (S - \overline{S}) \\ & + k_{as} \cdot (A - \overline{A}) \cdot (S - \overline{S}) + k_{ad} \cdot D \cdot (A - \overline{A}) \end{aligned}$$

Table 3: Warfarin Modeling Comparison

Method	Warfarin Best Model Test MSE
Existing Warfarin PK	0.646
D3-white-box	**0.39**
D3-hybrid	**0.271**

New Discovered PK Warfarin Model: Expert commentary

- **Prof. Jean-Baptiste Woillard, Pharmacologist.** "The **model is promising** and **pharmacokinetically plausible**. The next step is to apply D3 to other clinically relevant PK drug datasets."
- **Prof. Richard Peck, Clinical Pharmacologist.** "This model is reasonable and potentially superior. It represents a **significant advance in clinical pharmacology** by automatically identifying robust PK models."
- **Prof. Eoin McKinney, Clinician.** "**This model is significant**, as consortiums are dedicated to improving Warfarin [Consortium, 2009]. The **model adds novel components**, such as the Michaelis component for time-varying changes and novel interaction terms like age-sex."

Will LLM-Agents replace researchers?

Discovery of governing equations using ML

	Explicit function	Implicit function	Ordinary differential equation	Partial differential equation
Typical form	$y = f(x, t)$	$f(x, y) = c$	$\frac{dx}{dt} = f(x, t)$	$\frac{\partial u}{\partial t} = f(u, x)$

A SUPER hard problem

Why do we care about PDEs?

While **ODEs** model **time-dependent** processes in medicine, PDEs are necessary when **spatial variations** are important

Medical System	Requires PDEs?	Why?
Drug diffusion in tissues	✅ Yes	Spatial transport of drugs through organs/tissues.
Cancer growth & metastasis	✅ Yes	Tumors spread non-uniformly in space.
Cardiac & neural conduction	✅ Yes	Wave propagation in heart and brain.
Blood flow (hemodynamics)	✅ Yes	Navier-Stokes equations describe fluid dynamics.
Oxygen transport in tissues	✅ Yes	Oxygen diffuses through capillaries to tissues.
Tissue deformation & biomechanics	✅ Yes	Soft tissue and bone experience mechanical forces.
Epidemic spread with spatial effects	✅ Yes	Diseases spread differently across regions.
Radiation therapy dose planning	✅ Yes	Modeling energy transport through tissues.

What about higher order ODEs and PDEs?

$$\frac{du}{dt}$$

$$\frac{\partial u}{\partial t} \quad \frac{\partial u}{\partial x} \quad \frac{\partial u}{\partial y}$$

$$\frac{\partial^2 u}{\partial t^2} \quad \frac{\partial^2 u}{\partial x^2} \quad \frac{\partial^2 u}{\partial y^2}$$

$$\frac{\partial^2 u}{\partial t \partial x} \quad \frac{\partial^2 u}{\partial t \partial y} \quad \frac{\partial^2 u}{\partial x \partial y}$$

$$u\frac{\partial u}{\partial t} \quad u^2\frac{\partial u}{\partial t} \quad u\frac{\partial u}{\partial x}$$

Difficult to search

Variational trick may not work

Kacprzyk, K., Qian, Z. & vdS
D-CIPHER: Discovery of Closed-form Partial Differential Equations (NeurIPS 2023)

Krzysztof Kacprzyk

Zhaozhi Qian

van_der_Schaar \ LAB

Any PDE: Derivative-bound and derivative-free part

$$\underbrace{f(\boldsymbol{x}, \boldsymbol{u}(\boldsymbol{x}), \partial^{[K]}\boldsymbol{u}(\boldsymbol{x}))}_{\text{derivative-bound}} - \underbrace{g(\boldsymbol{x}, \boldsymbol{u}(\boldsymbol{x}))}_{\text{derivative-free}} = 0$$

Requires some technical constraints for the variational trick to work!

Can be evaluated directly from data – No additional constraints!

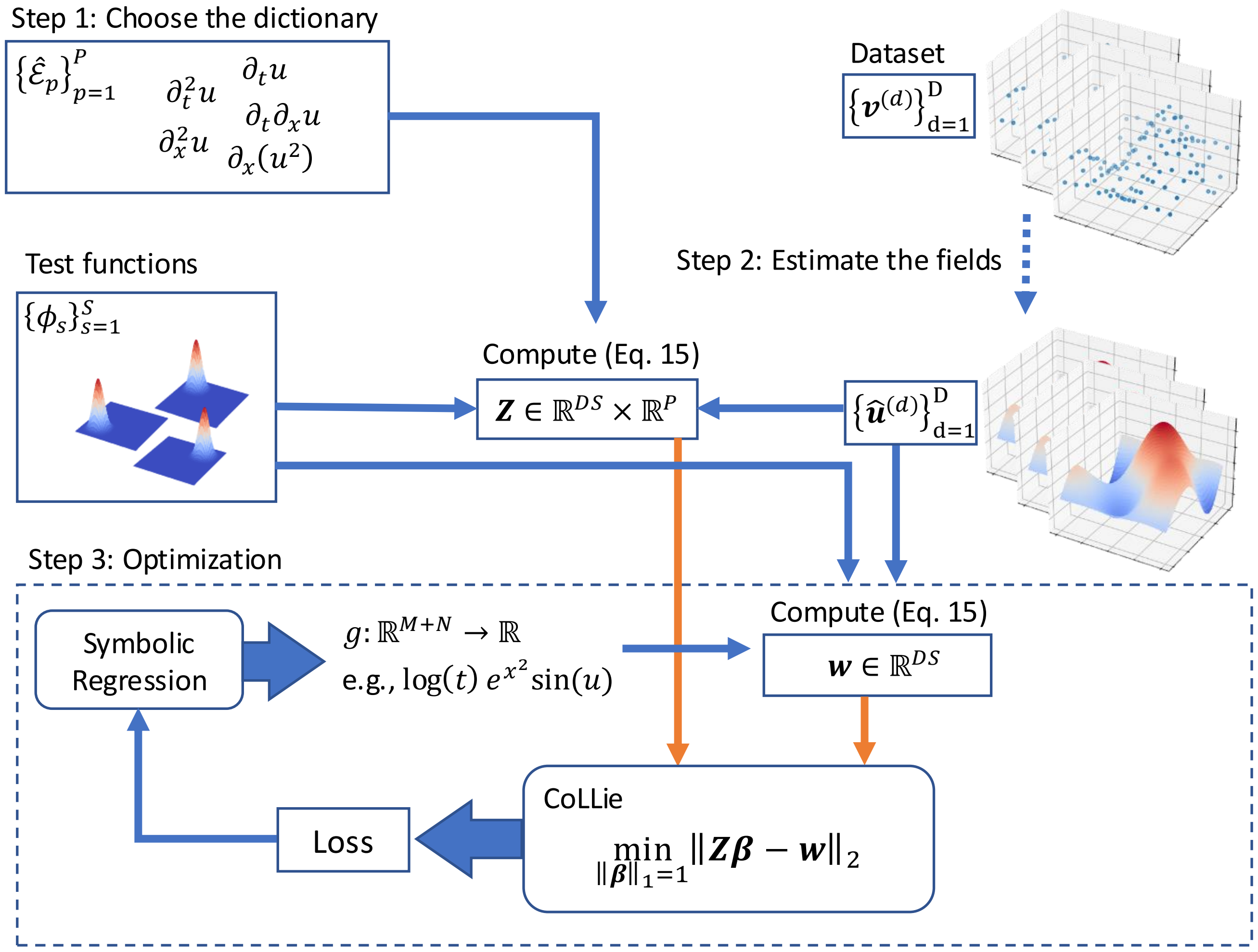
Variational-Ready PDEs

Currently the broadest family of PDEs that admit variational formulation

Kacprzyk, K., Qian, Z. & van der Schaar, M. D-CIPHER: Discovery of Closed-form Partial Differential Equations. (NeurIPS 2023)

D-CIPHER

- Algorithm
  - Uses variational formulation
  - Searches through all closed-form derivative-free parts
  - Searches through a linear subspace of derivative-bound parts



Do we need Equations to Understand Dynamical Systems?

Kacprzyk, K., & van der Schaar, M. (2025).
No Equations Needed: Learning System Dynamics Without Relying on Closed-Form ODEs.
***ICLR 2025*.**

Krzysztof Kacprzyk

Do we need Equations? Why Equations?

Syntactic representation

$$\frac{d\rho}{dt} = (\alpha_0 - \mu + \mu \cdot \Lambda) \cdot \rho - \mu \cdot \Lambda \cdot \rho^2$$

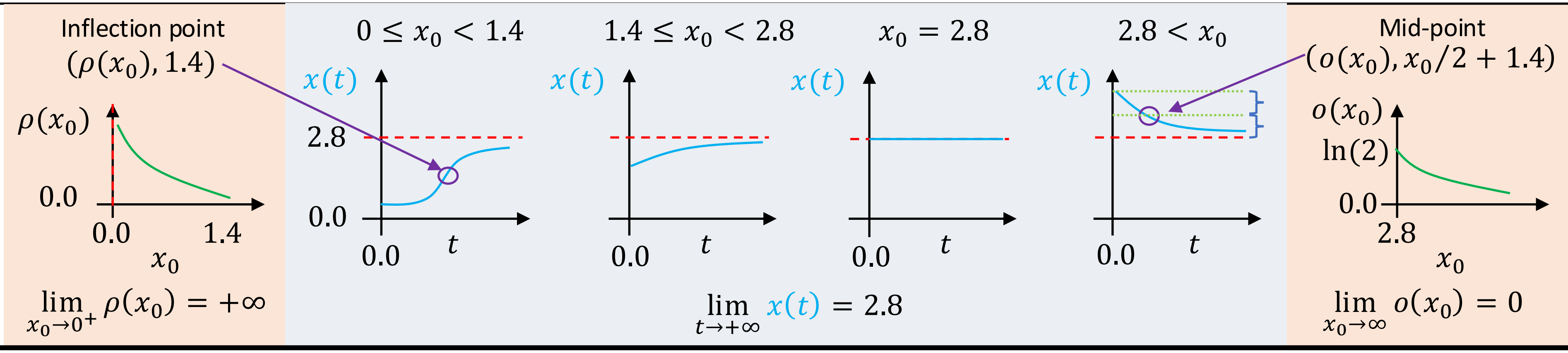
Logistic growth model

analysis

system behavior

Semantic representation

Syntactic vs. Semantic representation of an ODE

Syntactic representation	Semantic representation
$$\dot{x}(t) = x(t)\left(1 - \frac{x(t)}{2.8}\right)$$ $$x(0) = x_0$$	

- symbolic form

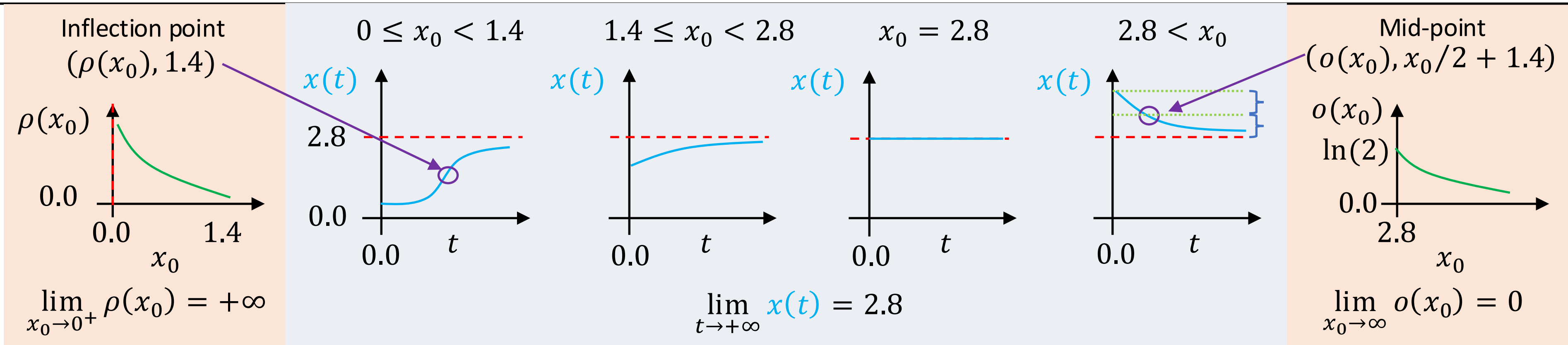
Semantic representation:

behaviour of a dynamical system

- shape, properties, asymptotic behavior
- how they change under different conditions

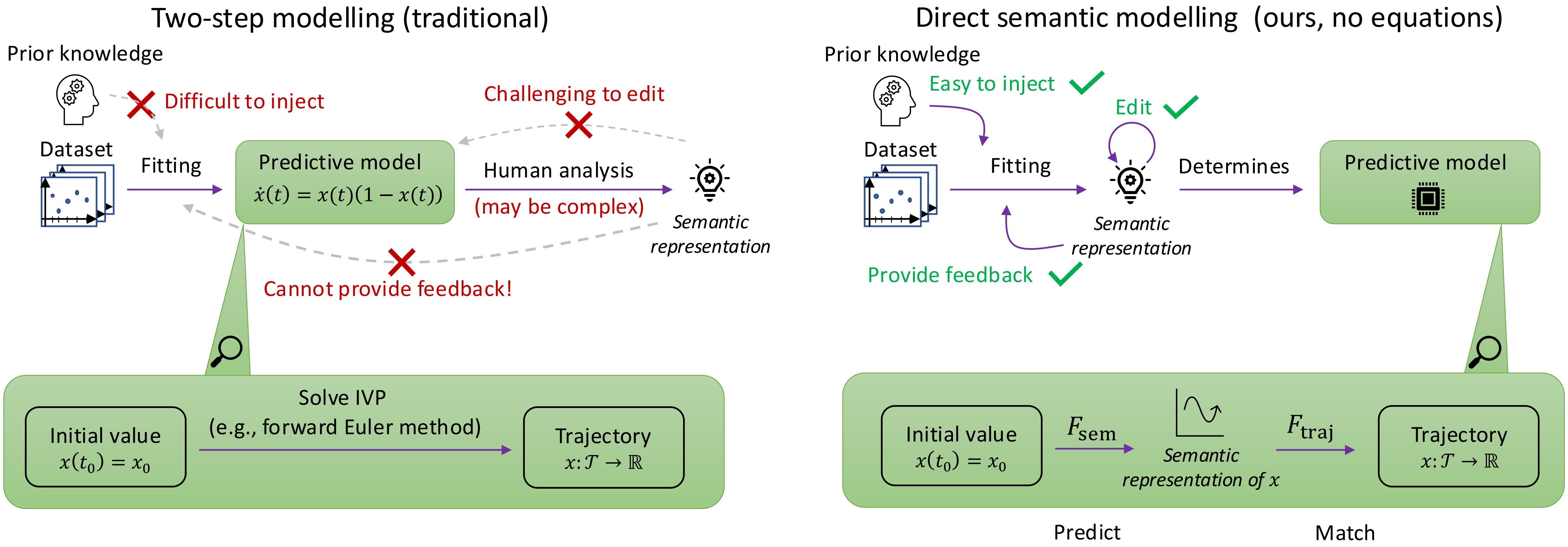
Semantic representation of a forecasting model, including a system of ODEs, describes changes under different initial cdts

Syntactic vs. Semantic representation of an ODE

Syntactic representation	Semantic representation
$\dot{x}(t) = x(t)\left(1 - \frac{x(t)}{2.8}\right)$, $x(0) = x_0$	

The focus is often on discovering equations …
and yet, what we actually want is to unravel these semantic representations

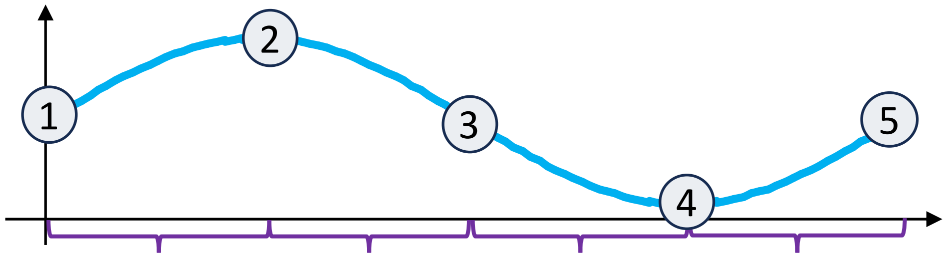
From Discovery and Analysis to Direct Semantic Modeling



What do we mean by semantic representation?

Formalizing Semantic Representations: Shape of trajectory

Kacprzyk, Liu, vdS, Towards Transparent Time Series Forecasting. ***ICLR 2024***



Motifs: s_{+-b} increasing concave | s_{--b} decreasing concave | s_{-+b} decreasing convex | s_{++b} increasing convex

Composition: $(s_{+-b}, s_{--b}, s_{-+b}, s_{++b})$

Transition points: 1 start | 2 local maximum | 3 Inflection point | 4 local minimum | 5 end

Krzysztof Kacprzyk

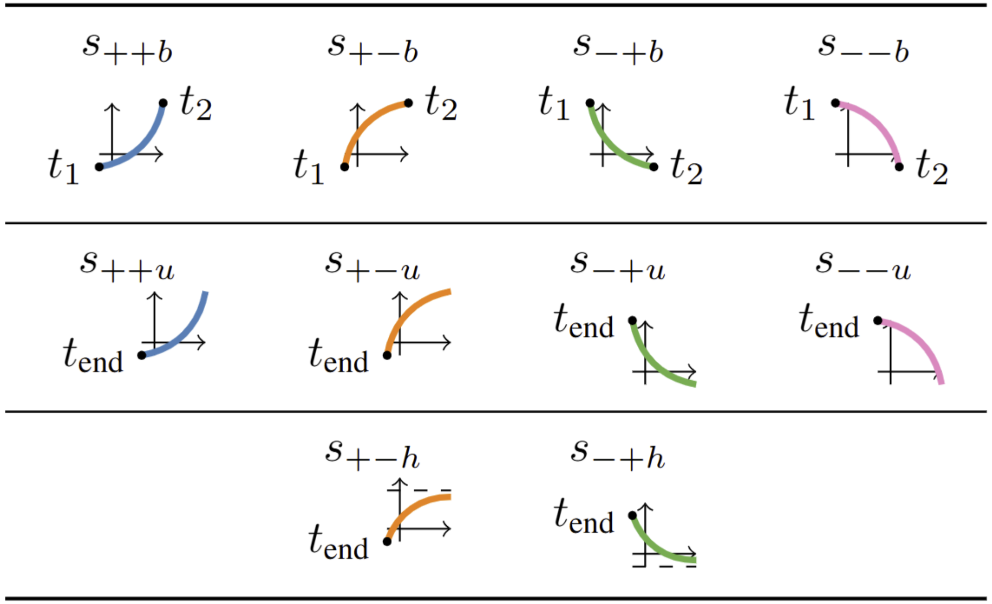
Divide the shape of trajectory into smaller intervals where the trajectory has a particular shape, called motifs

Formalizing Semantic Representations: Shape of trajectory

Kacprzyk, Liu, vdS, Towards Transparent Time Series Forecasting. *ICLR 2024*

Krzysztof Kacprzyk

In semantic ODEs, we extended this framework to add motifs that describe unbounded trajectories

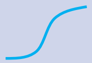


Formalizing Semantic Representations: Recap

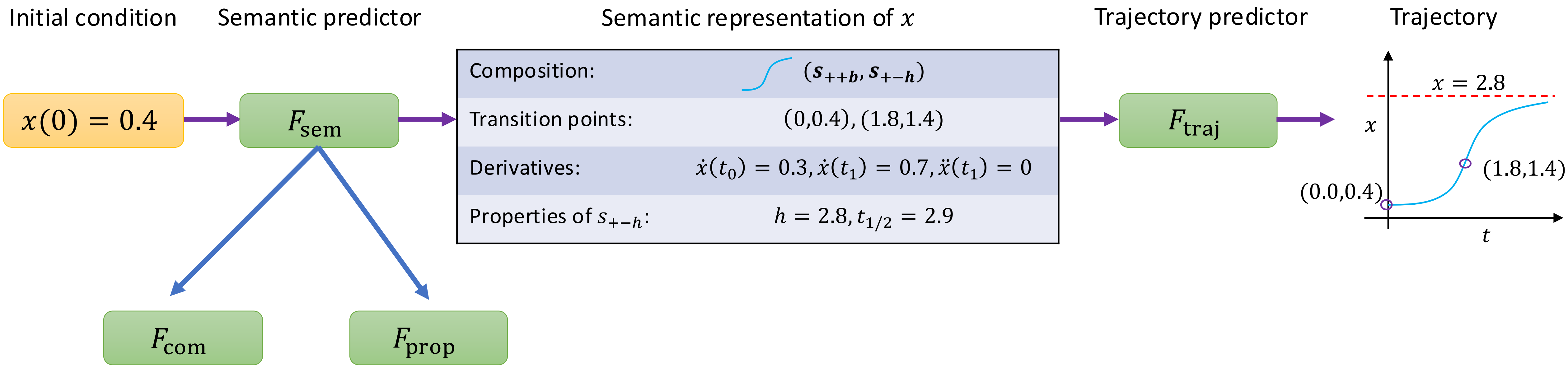
Semantic representation of x

c_x

p_x

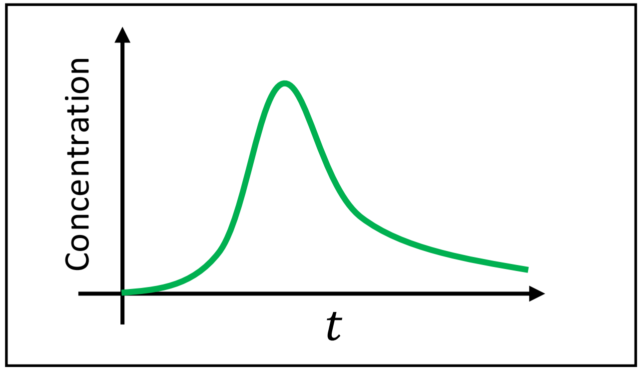
Composition:	 $(\boldsymbol{s}_{++b}, \boldsymbol{s}_{+-h})$
Transition points:	$(0,0.4), (1.8,1.4)$
Derivatives:	$\dot{x}(t_0) = 0.3, \dot{x}(t_1) = 0.7, \ddot{x}(t_1) = 0$
Properties of s_{+-h}:	$h = 2.8, t_{1/2} = 2.9$

How do we build Semantic ODEs?



Composition:	$(\boldsymbol{s}_{++b}, \boldsymbol{s}_{+-h})$
Transition points:	$(0{,}0.4), (1.8{,}1.4)$
Derivatives:	$\dot{x}(t_0) = 0.3, \dot{x}(t_1) = 0.7, \ddot{x}(t_1) = 0$
Properties of s_{+-h}:	$h = 2.8, t_{1/2} = 2.9$

Semantic Inductive Biases



ODE discovery

$$\dot{x}(t) = \sum_{i=1}^{n} \alpha_i g_i(t)$$

$n \leq 5$

$n \leq 10$

xt

x^3 x^2

$\log(t)$

Semantic ODE

Semantic motifs

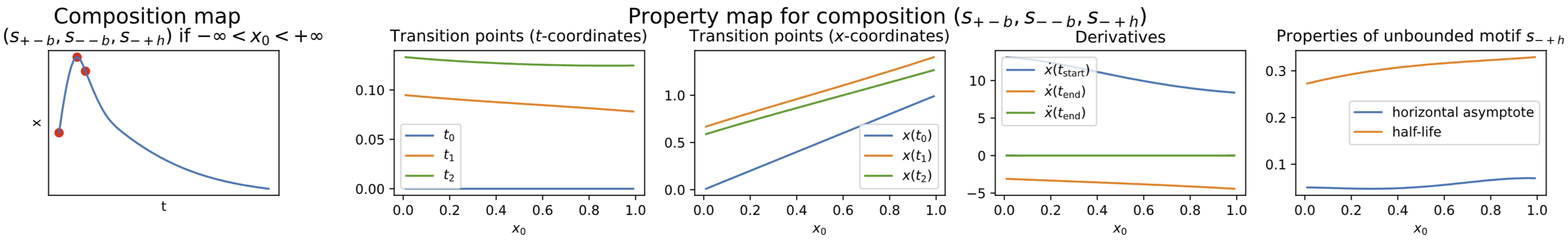
(s_{-+h}) (s_{--b}, s_{-+h}) $(s_{+-b}, s_{--b}, s_{-+h})$

$(s_{-+b}, s_{--b}, s_{-+h})$ $(s_{++b}, s_{+-b}, s_{--b}, s_{-+h})$

Can select number and types of motifs

Comprehensibility

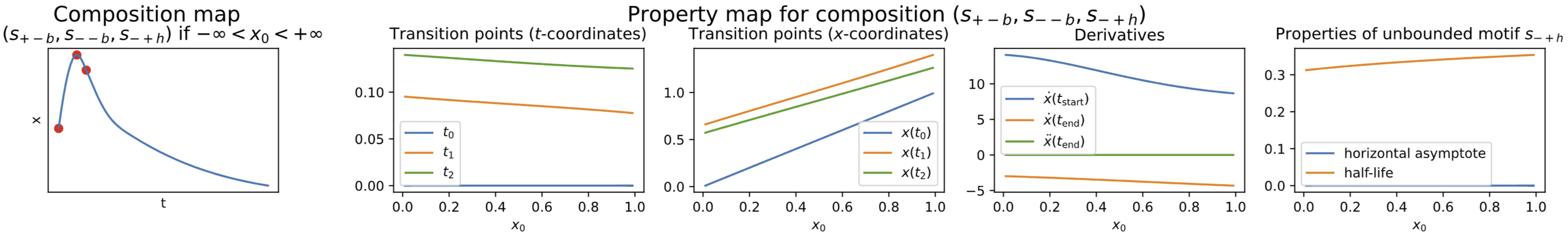
Model	Syntactic biases	Semantic biases	Syntactic representation	Semantic representation	In-domain ($t \leq 1$)	Out-domain ($t > 1$)
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 1$	NA	$\dot{x}(t) = -3.06x(t)t$	NA	$0.222_{(0.041)}$	$0.024_{(0.005)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 2$	NA	$\dot{x}(t) = 5.56 - 43.10x(t)t$	NA	$0.112_{(0.027)}$	$0.054_{(0.010)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 5$	NA	Equation (7)	NA	$0.101_{(0.023)}$	$16.850_{(0.021)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 10$	NA	Equation (8)	NA	$0.029_{(0.005)}$	$18.686_{(0.003)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 15$	NA	Equation (9)	NA	$0.020_{(0.004)}$	$77.577_{(1.249)}$
Semantic ODE	NA	$\lvert c_x \rvert \leq 4$, c_x ends with s_{-+h}	NA	Figure 7	$\mathbf{0.016}_{(.004)}$	$0.033_{(.006)}$



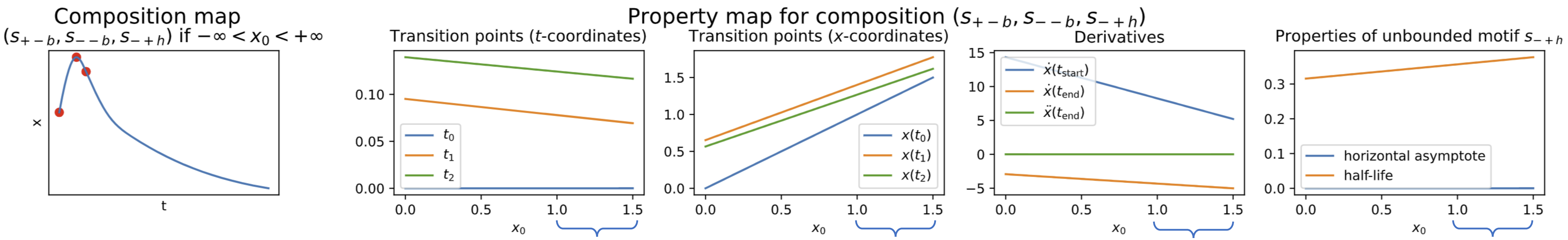
Editing

Edit and retrain our model while fixing this asymptote to 0

Model	Syntactic biases	Semantic biases	Syntactic representation	Semantic representation	In-domain ($t \leq 1$)	Out-domain ($t > 1$)
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 1$	NA	$\dot{x}(t) = -3.06x(t)t$	NA	$0.222_{(0.041)}$	$0.024_{(0.005)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 2$	NA	$\dot{x}(t) = 5.56 - 43.10x(t)t$	NA	$0.112_{(0.027)}$	$0.054_{(0.010)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 5$	NA	Equation (7)	NA	$0.101_{(0.023)}$	$16.850_{(0.021)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 10$	NA	Equation (8)	NA	$0.029_{(0.005)}$	$18.686_{(0.003)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 15$	NA	Equation (9)	NA	$0.020_{(0.004)}$	$77.577_{(1.249)}$
Semantic ODE	NA	$\lvert c_x \rvert \leq 4$, c_x ends with s_{-+h}	NA	Figure 7	$\mathbf{0.016}_{(.004)}$	$0.033_{(.006)}$



Extrapolation



Model	Syntactic biases	Semantic biases	Syntactic representation	Semantic representation	$x_0 \in (0, 1)$	$x_0 \in (1, 1.5)$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 1$	NA	$\dot{x}(t) = -3.06x(t)t$	NA	$0.222_{(0.041)}$	$0.240_{(0.035)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 2$	NA	$\dot{x}(t) = 5.56 - 43.10x(t)t$	NA	$0.112_{(0.027)}$	$0.131_{(0.025)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 5$	NA	Equation (7)	NA	$0.101_{(0.023)}$	$7.764_{(4.938)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 10$	NA	Equation (8)	NA	$0.029_{(0.005)}$	$0.105_{(0.056)}$
SINDy	$\dot{x} = \sum_{i=1}^{n} \alpha_i g_i, n \leq 15$	NA	Equation (9)	NA	$0.020_{(0.004)}$	$0.203_{(0.430)}$
Semantic ODE	NA	$c_x : (s_{+-b}, s_{--b}, s_{-+h}), h = 0$	NA	Figure 13	$\mathbf{0.018}_{(.003)}$	$\mathbf{0.023}_{(.005)}$

Great performance without seeing without a single sample from this distribution

Extrapolation for different inputs, outside training domain

van_der_Schaar \LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

Flexibility

Not only closed formed ODEs!

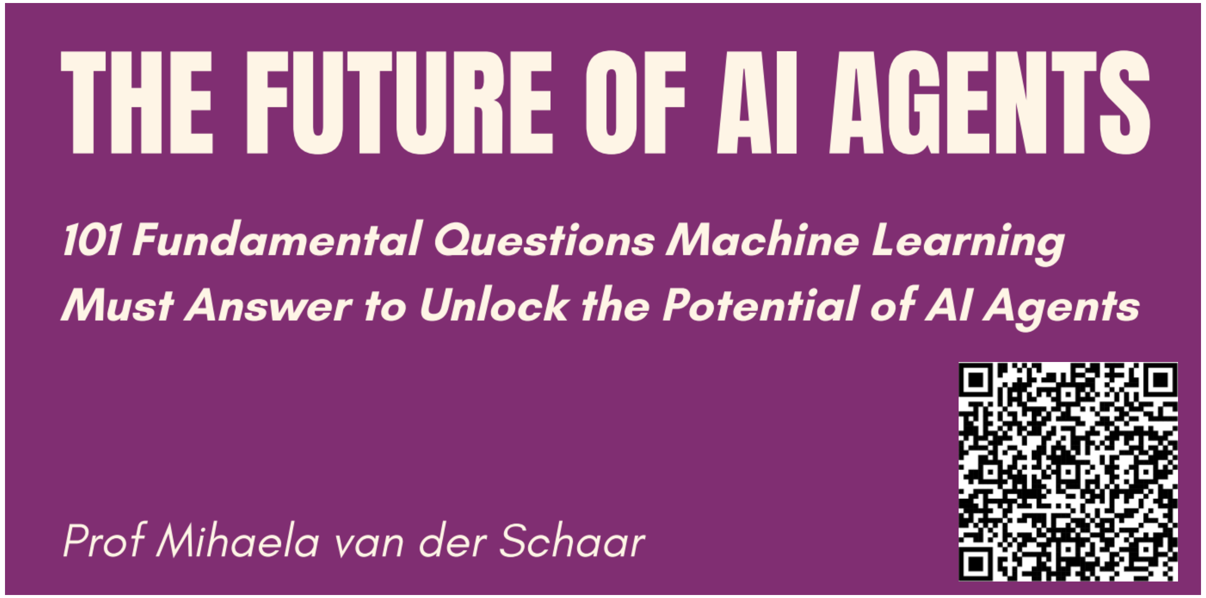
Method	Logistic Growth low noise	Logistic Growth high noise	General ODE low noise	General ODE high noise	Pharmacokinetic model low noise	Pharmacokinetic model high noise	Mackey-Glass (DDE) low noise	Mackey-Glass (DDE) high noise	Integro-DE low noise	Integro-DE high noise
SINDy-5	$0.012_{(.002)}$	$0.222_{(.004)}$	$0.053_{(.012)}$	$0.103_{(.010)}$	$0.093_{(.004)}$	$0.230_{(.014)}$	$0.238_{(.023)}$	$0.248_{(.025)}$	$0.431_{(.051)}$	$0.268_{(.019)}$
WSINDy-5	$\mathbf{0.010}_{(.000)}$	$0.222_{(.009)}$	$0.066_{(.009)}$	$0.102_{(.008)}$	$0.211_{(.009)}$	$0.415_{(.299)}$	$0.272_{(.032)}$	$0.300_{(.061)}$	$0.160_{(.066)}$	$0.452_{(.365)}$
PySR-20	$0.012_{(.002)}$	$0.224_{(.007)}$	$0.078_{(.029)}$	$0.119_{(.029)}$	$0.053_{(.015)}$	$0.242_{(.039)}$	$0.261_{(.021)}$	$0.288_{(.031)}$	$0.027_{(.011)}$	$0.393_{(.144)}$
SINDy	$0.012_{(.001)}$	$0.218_{(.011)}$	$0.068_{(.013)}$	$0.115_{(.012)}$	$0.020_{(.001)}$	$0.209_{(.010)}$	$0.252_{(.026)}$	$0.257_{(.028)}$	$0.318_{(.172)}$	$0.248_{(.016)}$
WSINDy	$\mathbf{0.010}_{(.001)}$	$0.217_{(.016)}$	$0.062_{(.009)}$	$0.112_{(.009)}$	$0.038_{(.006)}$	$0.219_{(.016)}$	$0.200_{(.035)}$	$0.207_{(.031)}$	$0.152_{(.086)}$	$0.300_{(.082)}$

This talk

We are entering a new **paradigm of discovery,** driven not by static models or datasets, but by **AI agents** capable of reasoning, experimentation, and evolving knowledge, transforming science from human-centric to **co-creative**

- LLM-based AI Agents
  - Why are they exciting
  - Examples
- Genies: Next-generation AI Agents

Genies – The Future of AI Agents



https://www.vanderschaar-lab.com/genies-the-future-of-ai-agents/

Introducing Genies: Next-generation AI agents

- What Are Genies?
  - Sophisticated, multi-capable AI “companions” that generate ideas, plan strategies, reason adaptively, and learn continuously.
  - Designed to amplify human potential rather than replace it.
- Goal
  - Empower humans through collaboration, trust, and transparent decision-making.
  - Move beyond narrow task automation or purely reactive “agents.”

Genies: The Synergy Loops

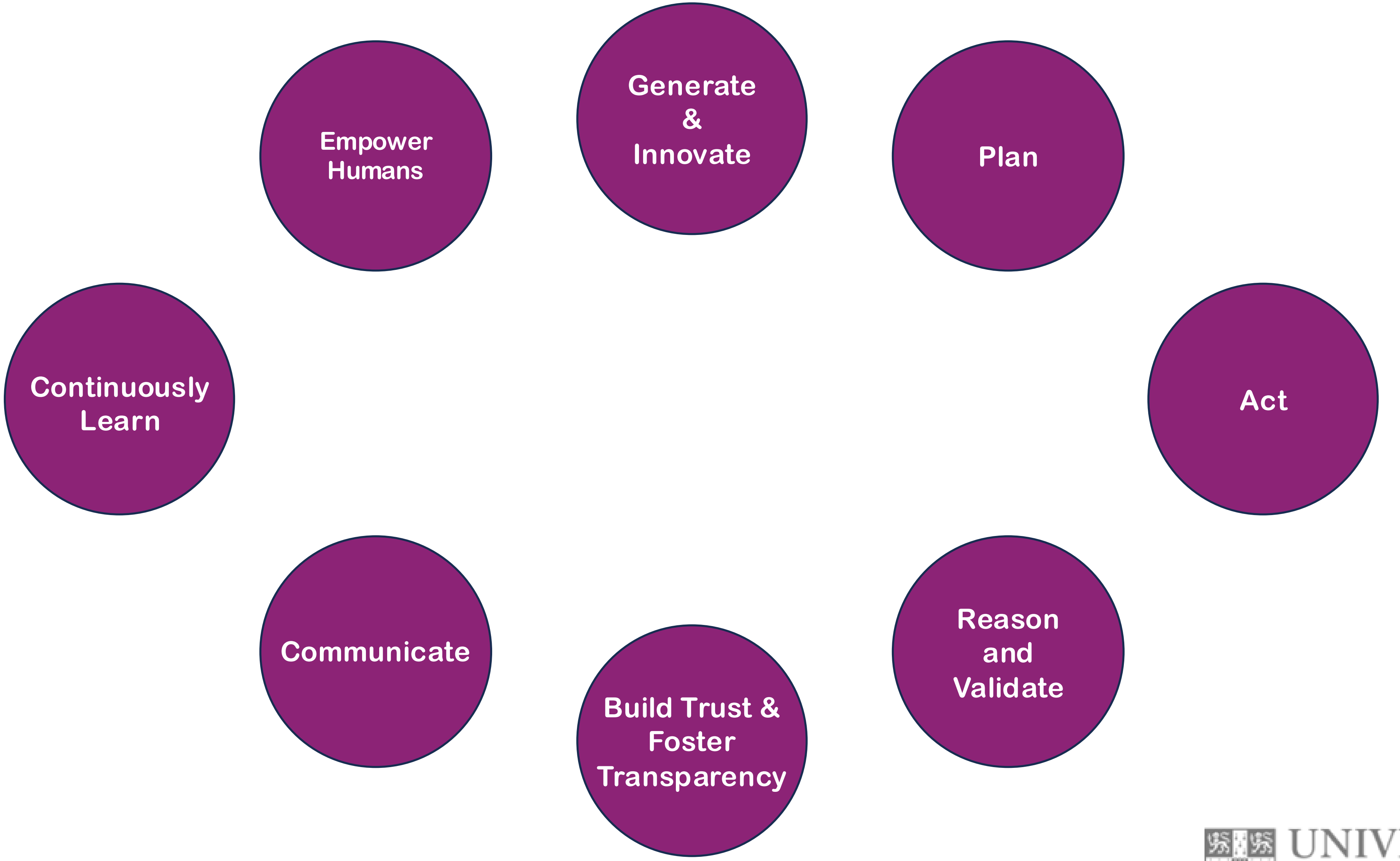
Five Interconnected Steps

- Innovation: Generating novel ideas, challenging assumptions.
- Operation: Planning and executing complex strategies.
- Validation: Rigorous reasoning and testing outcomes.
- Evolution: Continuous communication and learning.
- Empowerment: Building trust, expanding human abilities.

Self-Improving Cycle

- Each step reinforces the others, allowing genies to adapt over time.
- Designed for long-term partnership with humans (and other genies).

Genies – Core Capabilities



LLM Agents vs Genies

Aspect	LLM Agent	Genie
Primary Focus	Language-driven task execution	**Multi-capable** synergy (ideas, plans, validation, etc.)
Adaptation	Often short-term or prompt-based	**Lifelong**, evolving with user & environment
Memory	Limited context window or ephemeral tools	**Persistent, structured, continuously integrated** knowledge
Innovation	Usually reliant on user prompt or fine-tunes	**Proactively** generates novel strategies & creative solutions
Human Role	User queries, agent responds	**Collaborative** co-creation and empowerment

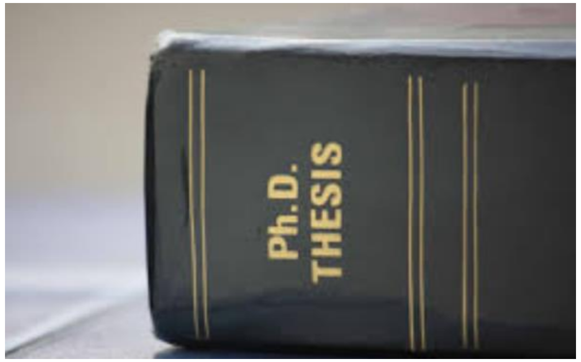
Genies aim for **comprehensive, long-term** synergy with humans, not just short bursts of text-based help.

A challenging agenda

101 Fundamental Questions Machine Learning Must Answer to Unlock the Potential of AI Agents

by Prof Mihaela van der Schaar

Answering each question is a



Why We Should Be Excited: Shaping the Future of AI + Human Collaboration

Transformative Partnerships

- Genies expand what’s possible: advanced healthcare support, scientific discovery, personalized education, etc.

Continuous Improvement

- Self-evolving loop ensures ongoing value rather than “one-and-done” or static solutions.

Potential Impacts

- Research & Innovation: New breakthroughs via cross-domain synergy.
- Everyday Life: Personal genies that evolve with your career, health, and creative pursuits.

This talk

We are entering a new **paradigm of discovery,** driven not by static models or datasets, but by **AI agents** capable of reasoning, experimentation, and evolving knowledge, transforming science from human-centric to **co-creative**

- LLM-based AI Agents
  - Why are they exciting
  - Examples
- Genies: Next-generation AI Agents
- Agent Networks

The Agent Network

The Agent Network

Unleashing Self-Organising AI
for a New Economic Era

Prof Mihaela van der Schaar

https://www.vanderschaar-lab.com/the-agent-network/

Rethinking Scientific Discovery using AI Agent Networks


Tennison Liu


Silas Estevez


David Bentley


Mihaela
van der Schaar

Hypothesis Hunting

Hypothesis hunting is the continuous and diverse exploration of large-scale datasets to surface promising findings that guide subsequent human investigation and experimental validation.

Rethinking Scientific Discovery: Why Hypothesis Hunting Needs AI Agents

Modern science faces a **crisis of complexity**
- **Scale:** Millions of data points × thousands of variables = **combinatorial explosion**
- **Coordination:** Knowledge, tools, and perspectives scattered across disciplines
- **Human bottleneck:** Limited capacity for exhaustive exploration and synthesis

Scale + Coordination → Beyond Human Reach Alone

The Opportunity: Autonomous, evolving networks of AI agents can
– explore vast hypothesis spaces,
– iteratively refine candidate findings,
– accumulate and share knowledge,
surfacing **testable hypotheses** for human interpretation and experimental validation.

Beyond Today's "AI Scientists": Why We Need True Hypothesis-Hunting Agents

Recent systems like **AI Scientists** and **Co-Scientists** (Lu et al., 2024; Gottweis et al., 2025) represent important progress - can propose hypotheses, design experiments, run analyses, and interpret results

Limitations of Current AI Scientist Paradigm:

- **Predefined Questions** → operate within narrow, human-specified problem scopes
- **Limited Exploration** → insufficient coverage of vast, sparse hypothesis spaces
- **Weak Evaluation** → struggle to assess significance without a fixed question context
- **No Cumulative Knowledge** → focus on isolated outputs, not evolving, layered research programs

Beyond Today's "AI Scientists": Why We Need True Hypothesis-Hunting Agents

What Hypothesis Hunting Requires:

- Exploration – search vast, open-ended hypothesis spaces
- Evaluation – assess heterogeneous, context-dependent discoveries
- Accumulation – refine, layer, and recombine findings into evolving programs
- Evolution – adapt reasoning as knowledge grows

We need agents that don't just answer our questions — they help us find the questions worth asking

How can we unlock Hypothesis Hunting at Scale?

Insight: Scientific progress emerges from networks, not individuals.
To advance systematic discovery, we need **networks of autonomous AI agents**, not solitary ones.

In human science, breakthroughs arise from:

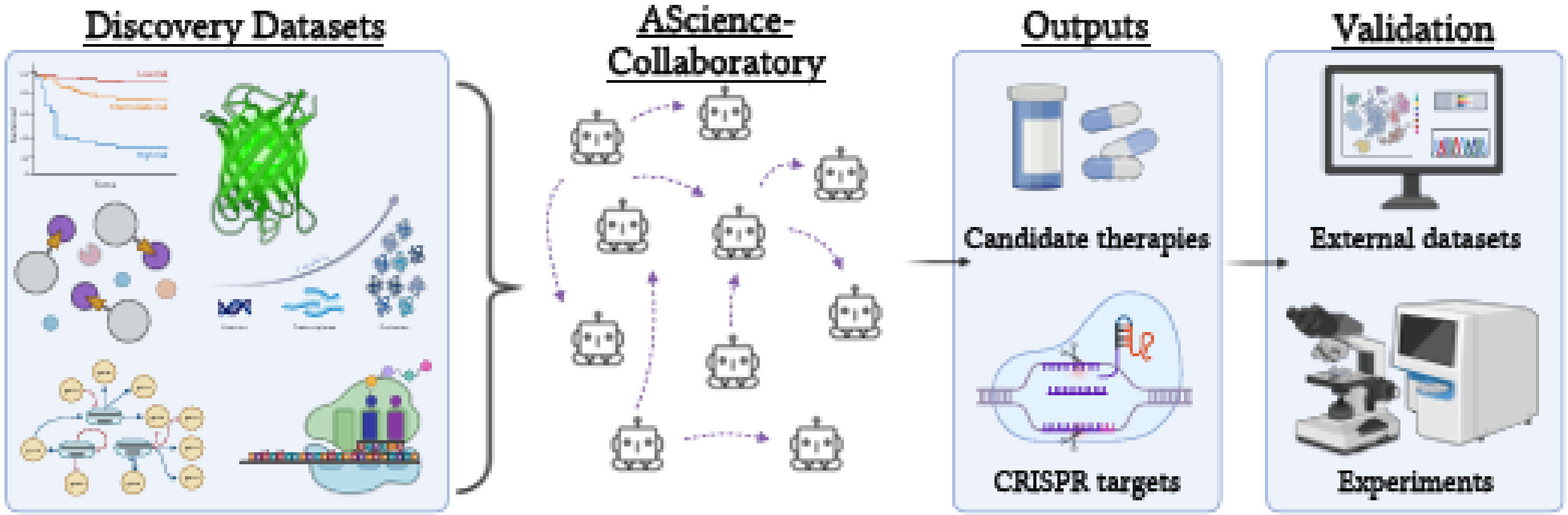
- **Diverse perspectives** across domains
- **Cross-pollination** of methods and ideas
- **Critical debate** and iterative refinement
- **Cumulative evidence** built across time

Their **social dynamics** — exploration diversity, critique, knowledge exchange — are **crucial** to uncover novel directions and accumulate layered understanding.

We aim to replicate and amplify these dynamics in AI agentic systems.

Hypothesis Hunting with Evolving Networks of Autonomous Scientific Agents

Large-scale datasets are explored by autonomous networks of research agents that collaborate, peer-review, and refine findings to surface promising directions for human validation



Our Framework: *AScience*

AScience formalizes **hypothesis hunting** as a **continuous, open-ended exploration** process, embedding the **social dynamics** of cumulative scientific progress.

Key elements:

- **Distributed agents** exploring diverse directions
- **Endogenous interaction** (sharing, critique, refinement)
- **Shared evaluation frameworks** for coherence
- **Iterative accumulation** of knowledge into evolving research programs

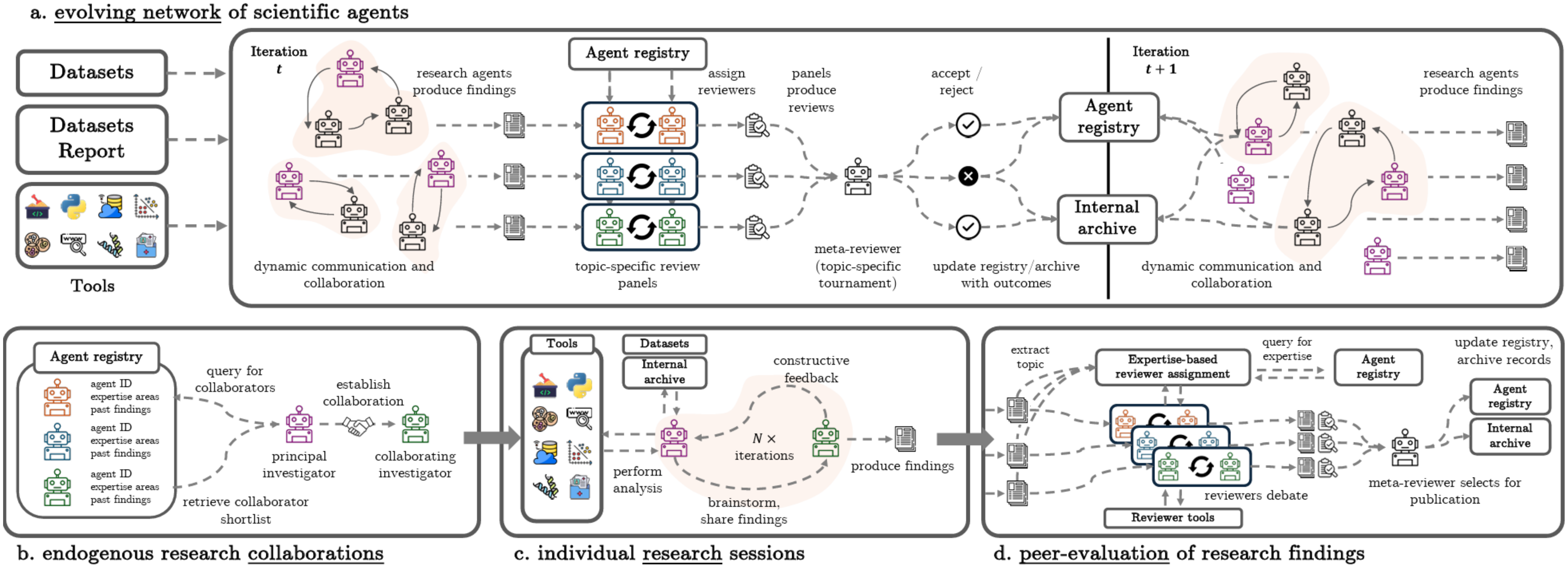
Social dynamics unlock collective intelligence in scientific discovery.

Implementation: *ASCollab*

AScience instantiated as **ASCollab**, a network of heterogeneous LLM-based research agents, that:

- Generate and test hypotheses
- Critique and refine one another’s findings
- Share evaluation standards
- Build cumulative knowledge collaboratively

ASCollab:
Paving the way for scalable, directed & open-ended hypothesis hunting



The Cancer Genome Atlas (TCGA): A Prime Testbed for Autonomous Hypothesis Discovery

TCGA offers an **ideal environment** to evaluate ASCollab's ability to discover novel, meaningful hypotheses from large-scale, real-world biomedical data.

Why TCGA?

1. Real-World Impact

- Understanding **new mechanisms**, **biomarkers**, and **therapeutic targets** in cancer remains a major open challenge.
- Hypotheses uncovered here can drive **scientific insight** and **clinical translation**.

2. Scale & Richness

- **Multi-omics** data across dozens of cancer types.
- Vast, sparsely explored **combinatorial space** of possible associations & mechanistic hypotheses.

3. Reproducibility

- **Open-access** resource ensures findings can be validated and compared by the community.

ASCollab Demonstrates Promise for Hypothesis Hunting at Scale

Can a ASCollab surface novel, interpretable, and biologically meaningful hypotheses (discoveries) from TCGA data?

Representative Discoveries found via ASCollab:

- **Rediscoveries** of **established cancer drivers** → *validation of the system's reliability and grounding in known biology*
- **Extensions of pathways** → *novel mechanistic insights building on emerging literature*
- **Proposals of new candidate therapeutic targets** → *previously unexplored associations with survival outcomes*

These preliminary results showcase how **networked scientific agents** can

- navigate high-dimensional biomedical data
- surface **interpretable**, **biologically plausible**, and **clinically relevant** hypotheses
- bridge **discovery and translational insight** at unprecedented scale

Multi-gene Ferroptosis axis in KIRC (Section D.1)

Agents identified a ferroptosis module involving ACSL4, GPX4, and FTH1 in kidney cancer, a part of which was later independently discovered and published in Zheng et al. (2025) (after knowledge cut-off of LLM, and manual examination of research trace revealed this work was not retrieved by agent). This finding, supported by DepMap essentiality data and prior mixed evidence (Guo et al., 2015; Huang et al., 2019; Zou et al., 2019), was enabled by the primary agent extending earlier findings by another agent (on SLC7A11/ALOX5) into a broader mechanistic hypothesis.

SLC5A2 and ABCC8 in PAAD (Section D.2)

Agents proposed SLC5A2 (SGLT2) and ABCC8 as therapeutic targets in pancreatic adenocarcinoma, anticipating a July 2025 publication that independently confirmed the SLC5A2−PAAD link (Xie et al., 2025). This finding, contextualized against prior work emphasizing SGLT1 (Du et al., 2022) and largely non-oncologic studies of SGLT2 (Jurczak et al., 2011), illustrates how agent collaboration surfaced a novel target class while situating results within the transporter literature.

BIRC5 validation and PRKD1 extension in KIRC (Section D.3)

Agents independently reproduced the established role of BIRC5 (Survivin) as a diagnostic and prognostic marker in KIRC (Wang et al., 2021), strengthening confidence by re-deriving results from scratch on TCGA data. Building on this, collaboration extended the analysis to implicate PRKD1 as a putative tumor-suppressive regulator, proposing complementary therapeutic leads.

A fundamentally new paradigm is needed for AI to orient it towards addressing the complexities of the real-world

Genies – The Future of AI Agents

THE FUTURE OF AI AGENTS

101 Fundamental Questions Machine Learning Must Answer to Unlock the Potential of AI Agents

Prof Mihaela van der Schaar

https://www.vanderschaar-lab.com/genies-the-future-of-ai-agents/

GENIES

Next-Generation AI Agents

PhD Opportunities at the van der Schaar Lab

We are looking for 4 fully funded PhDs to join our lab! Applications now OPEN!

About our PhD programme

- Based at the **University of Cambridge (DAMTP)**
- Work at the cutting edge in a world-leading lab
- Projects with a purpose; work that can change the world
- **32** papers accepted at four largest AI/ML conferences (NeurIPS, ICML, ICLR, AISTATS) in the past year

Apply through our website:
www.vanderschaar-lab.com/join-the-van-der-schaar-lab/

For any queries, please email:
vanderschaarlab@damtp.cam.ac.uk

Join our PhD Open Day

- **Monday 20 October, 9:30am (BST)** *(virtual)*
- Hear from current PhD students, explore research areas, and ask questions

Scan the QR code to register for the PhD Open Day:

Let’s shape the future together!

Mihaela van der Schaar, FRS

John Humphrey Plummer Professor of Machine Learning, Artificial Intelligence and Medicine, University of Cambridge
Director, Cambridge Center for AI in Medicine

van_der_Schaar \ LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

mv472@cam.ac.uk

@MihaelaVDS

linkedin.com/in/mihaela-van-der-schaar/