Quantitative Epistemology:
Conceiving a new human-machine partnership

Mihaela van der Schaar

John Humphrey Plummer Professor of
Machine Learning, Artificial Intelligence and Medicine, University of Cambridge

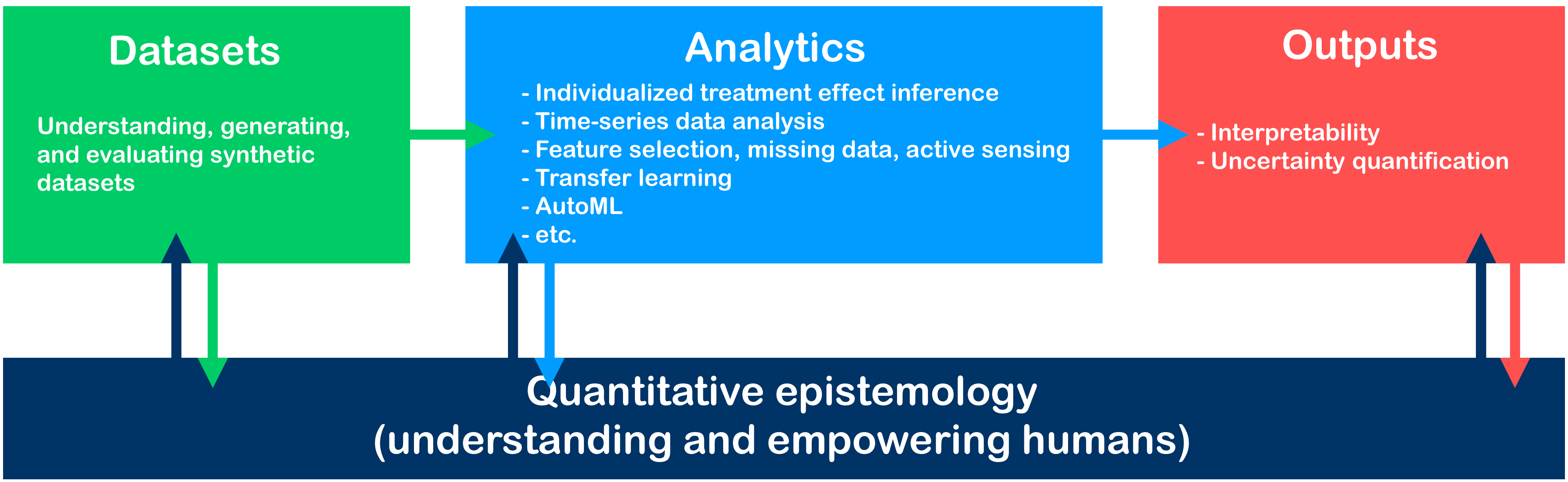
vanderschaar-lab.com

mv472@cam.ac.uk

@MihaelaVDS

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

Our group's research agenda:
New ML aimed at revolutionizing healthcare



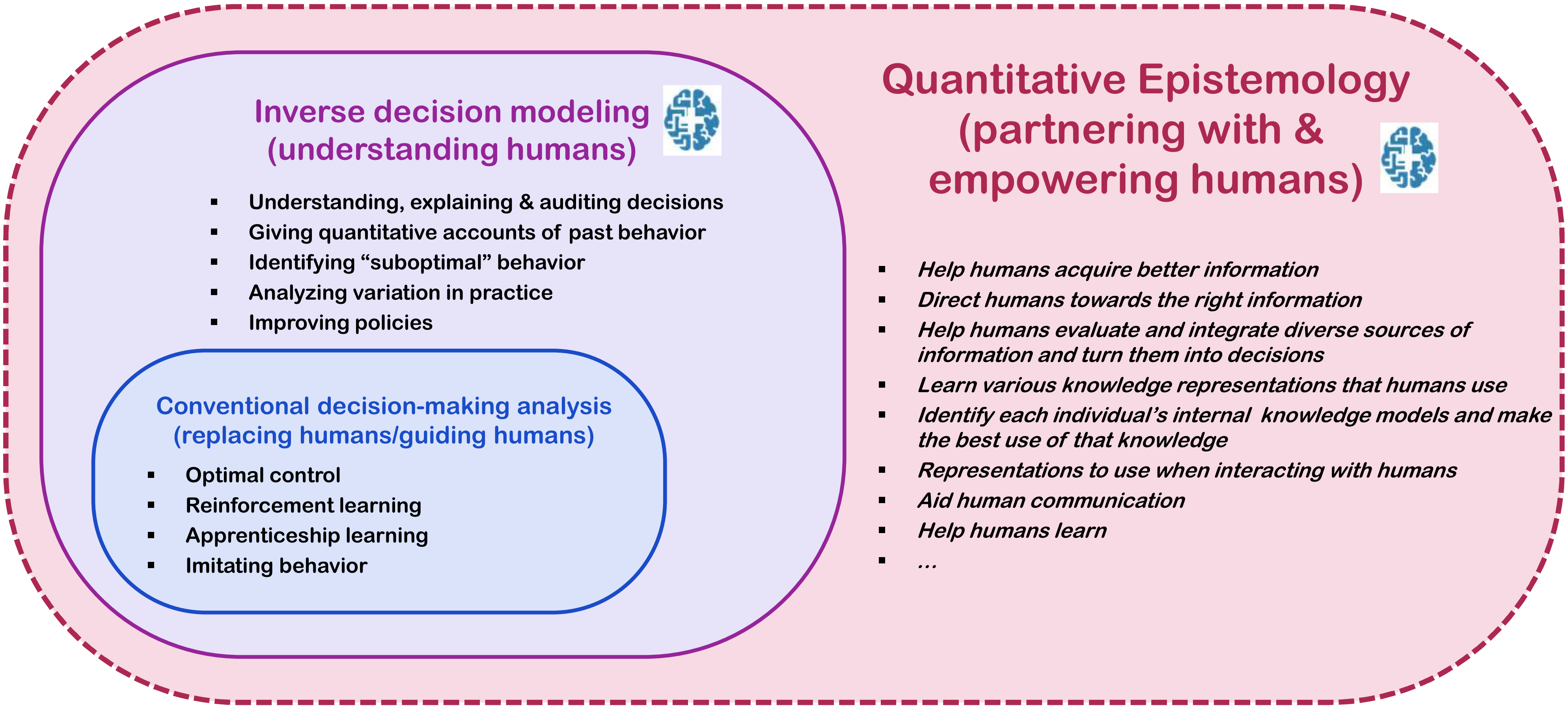
van_der_Schaar
\LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

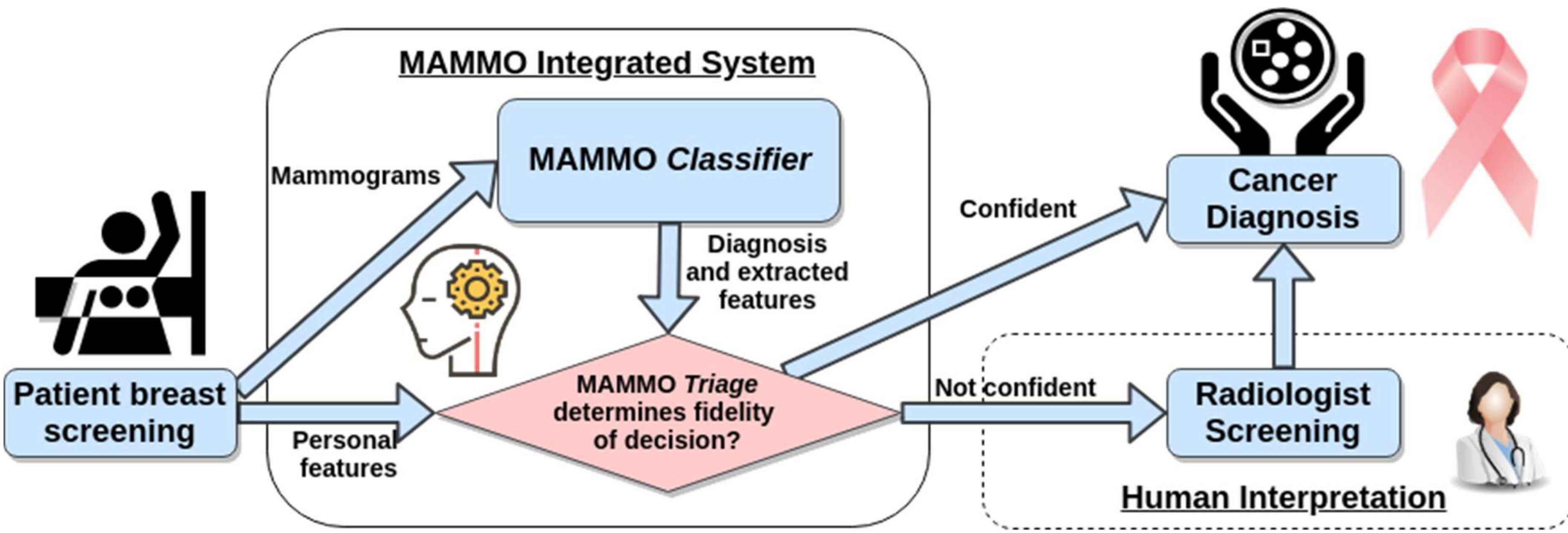
Explaining the name…

$$\underbrace{\text{Quantitative}}_{\text{Refers to things that can be measured}}\ \underbrace{\text{epistemology}}_{\text{The study of knowledge}}$$


Inverse decision modeling
(understanding humans)
Understanding, explaining & auditing decisions
Giving quantitative accounts of past behavior
Identifying “suboptimal” behavior
Analyzing variation in practice
Improving policies
Conventional decision-making analysis
(replacing humans/guiding humans)
Optimal control
Reinforcement learning
Apprenticeship learning
Imitating behavior
Quantitative Epistemology
(partnering with &
empowering humans)
Help humans acquire better information
Direct humans towards the right information
Help humans evaluate and integrate diverse sources of information and turn them into decisions
Learn various knowledge representations that humans use
Identify each individual’s internal knowledge models and make the best use of that knowledge
Representations to use when interacting with humans
Aid human communication
Help humans learn
...

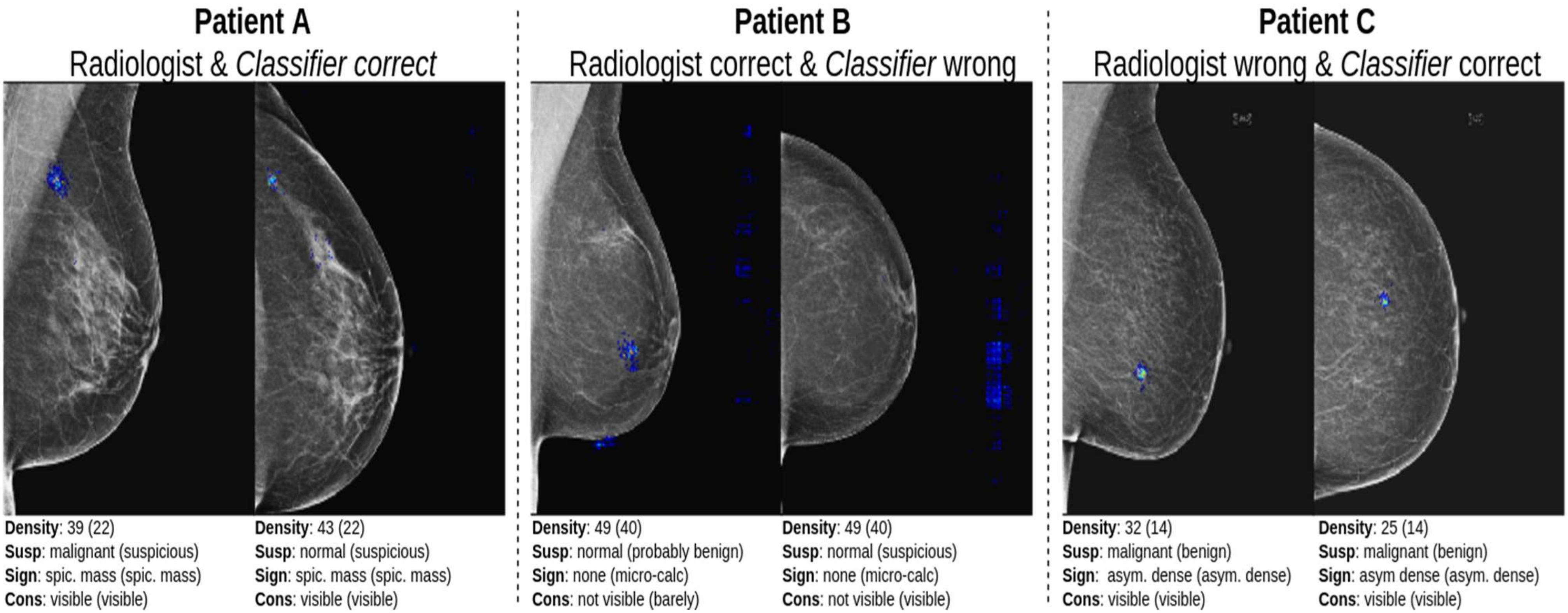
MAMMO: a framework for human-ML cooperation

[Kyono, vdS, ML4HC 2019]
[Kyono, Gilbert, vdS, JACR, 2019]



Machine learning for mammography article named "Best of 2020" by JACR

Who is better?
Human (radiologist) or machine (classifier)?



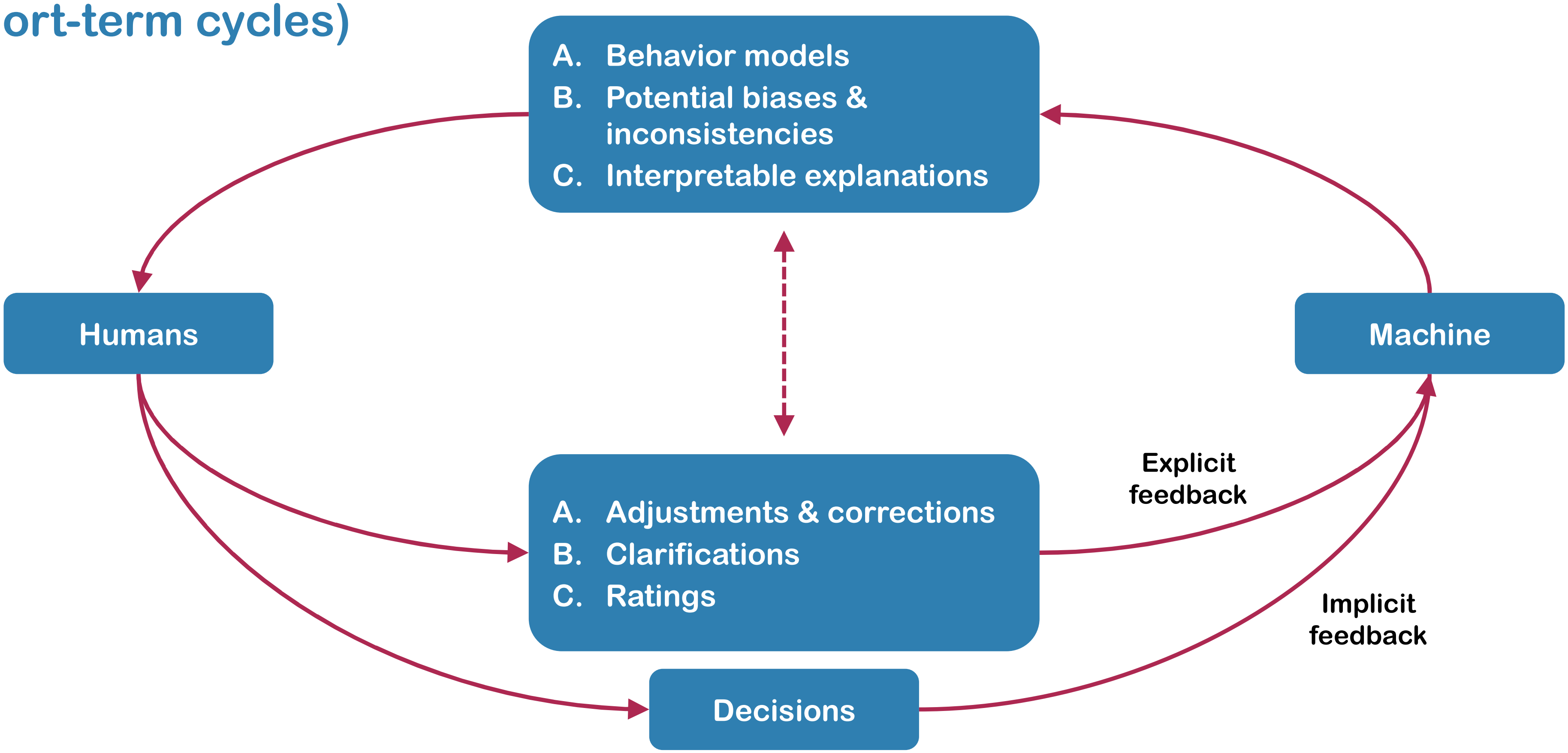
MAMMO enables various cooperation modes between humans and machines

MAMMO – Cooperation modes

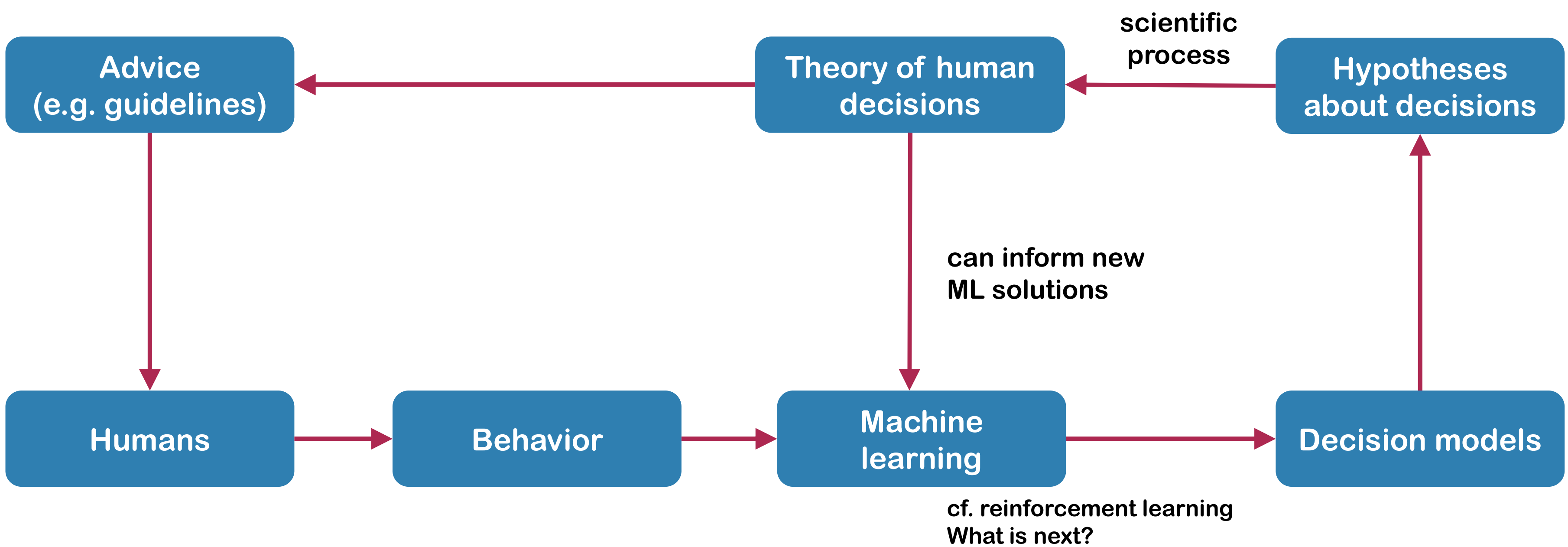
1. **Radiologist + Classifier – both activated as <u>double readers</u>**
2. **Radiologist + Classifier – triaging - operating as a <u>single reader</u> (hybridized)**

	Radiologist patients	*Classifier* patients	Cohen's κ	F1 score	TP	TN	FP	FN
Radiologist	1000	0	0.708	0.755	120	802	42	36
Classifier	0	1000	0.420	0.433	61	**811**	**33**	95
Classifier®	1000	1000	0.647	0.708	**125**	772	72	**31**
MAMMO	**456**	**544**	**0.724**	**0.766**	118	**810**	**34**	38

New human-machine partnership: Online operation
(short-term cycles)



New human-machine partnership (long-term cycles)



Quantitative Epistemology: Our work so far

IDM framework (ICML'21)

Agent = human

Method	Goal / motivating question	Planner	Normative params.	Descriptive params.
IAS (ICML'20)	How "timely" is agent decision making?	Timely active sensing	Deadline, cost of acquisition	Importance of accuracy, speed, efficiency
AVRIL (ICLR'21)	What reward function does the agent optimize?	RL planner	-	Reward function
CIRL (ICLR'21)	How important are various counterfactuals in making decisions?	Counterfactual RL planner	Counterfactuals	Importance weights
INTERPOLE (ICLR'21)	What are the subjective beliefs of the agent?	Policies based on decision boundaries	Interpretable state space	Decision dynamics & decision boundaries
IBRC (ICML'21)	How optimal is agent behavior relative to an "ideal" reward function?	Bounded rational planner	"Ideal" reward function	Flexibility, optimism, adaptivity
ICB (submitted)	How does agent's behavior evolve over time?	Contextual bandit strategies	-	Time-varying beliefs over reward functions

Inverse decision modeling (IDM) –
Learning Interpretable Representations of Behavior

Human decision-making is not perfect

- bounded rationality, cognitive biases

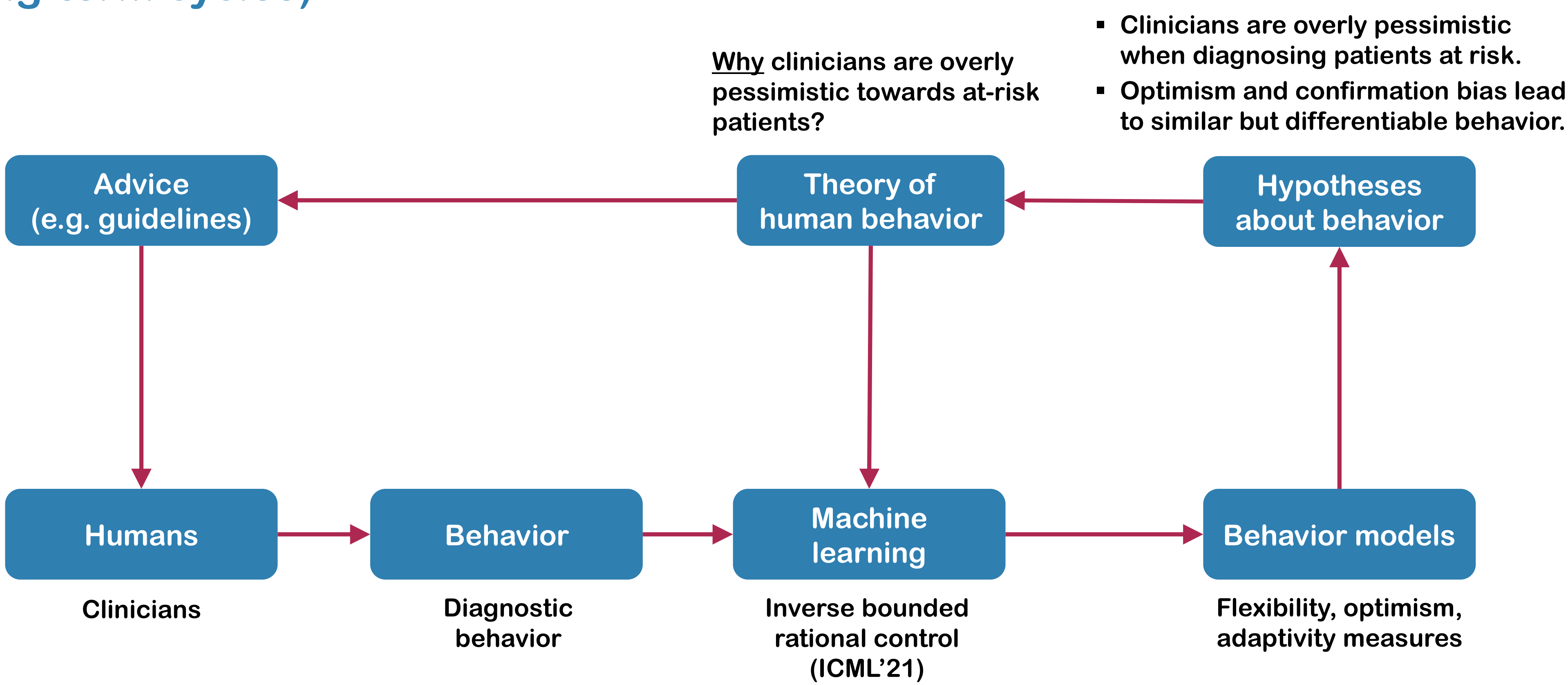
How can we help humans make better decisions?

- requires a quantitative account of the "imperfections" that necessitate correcting

Inverse decision modeling

- general framework for learning representations of decision-making behavior
- enables us to describe existing behavior relative to "ideal" behavior

New human-machine partnership within the IDM framework
(long-term cycles)



Conventional decision-making analysis

The "forward" problem:

- What constitutes ideal behavior?

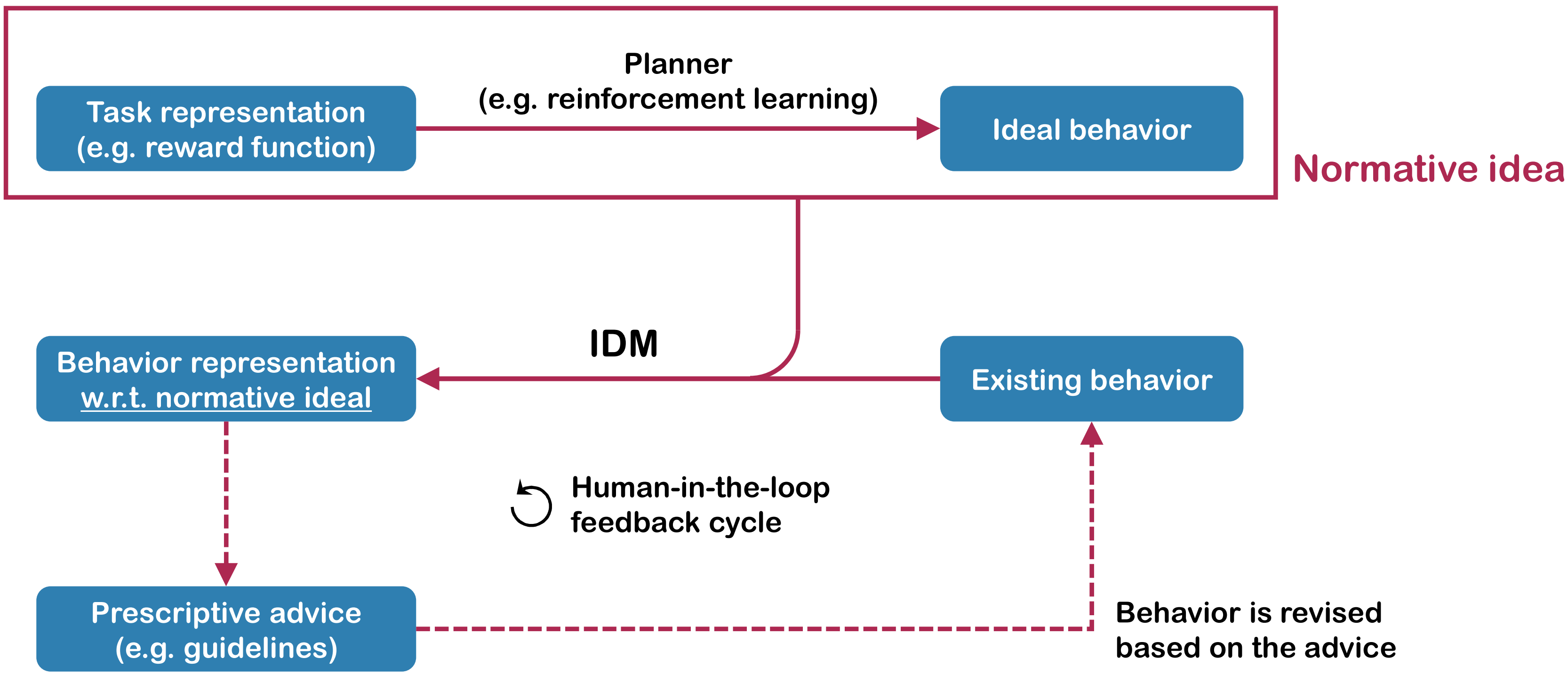
The "inverse" problem:

- What does the existing behavior look like?

Existing solutions offer limited help

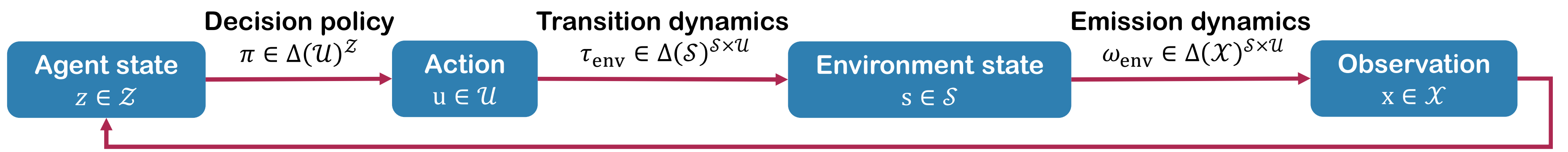
- forward solutions do not take human behavior into account
- inverse solutions focus on imitating human behavior

Inverse decision modeling



Planners

Problem setting:



- **Planning parameters:** $\theta \in \Theta$ **(e.g. utility function** $\upsilon \in \mathbb{R}^{\mathcal{S}\times\mathcal{U}}$**)**
- **Behavior:** $\phi_{\pi,\rho} \in \Phi = \Delta(\cup_t (\mathcal{X} \times \mathcal{U})^t)$

(Forward) planner:

$$F(\theta) = \phi_{\pi^*,\rho^*} \quad \text{where} \quad \pi^*, \rho^* = \text{argmax}_{\pi,\rho}\mathcal{F}(\pi,\rho;\theta)$$

- **e.g. reinforcement learning:** $\mathcal{F}(\pi;\theta) = \mathbb{E}_\pi[\sum_t \upsilon(s_t,u_t)]$

Inverse planners

- **Demonstrated behavior:** $\phi_{\text{demo}} \in \Phi$
- **Normative/descriptive params.:** $\theta = (\theta_{\text{norm}}, \theta_{\text{desc}}) \in \Theta = \Theta_{\text{norm}} \times \Theta_{\text{desc}}$

Inverse planner:

$$\hat{\theta}_{\text{desc}} = \text{argmax}_{\theta_{\text{desc}}} \mathcal{G}\big(\phi_{\text{demo}}, \phi_{\text{imit}} = F(\theta_{\text{norm}}, \theta_{\text{desc}})\big)$$

- **e.g. distribution matching:** $\mathcal{G}(\phi_{\text{demo}}, \phi_{\text{imit}}) = -D_{\text{KL}}(\phi_{\text{demo}} || \phi_{\text{imit}})$
- **projection of** ϕ_{demo} **onto** $\Phi_{\theta_{\text{norm}}} = F(\theta_{\text{norm}}, \Theta_{\text{desc}})$

- **Subsumes a wide range of algorithms**
- **Opens up new possibilities for interpretative research on decision making**

An application of IDM

Inverse reinforcement learning

- F = the RL planner
- $\theta_{\text{norm}} = \emptyset$
- $\theta_{\text{desc}} = \upsilon$ (reward/utility function)
- $\mathcal{G}\big(\phi_{\text{demo}}, \phi_{\text{imit}} = F(\upsilon)\big) = \mathbb{E}[V_\upsilon(\phi_{\text{demo}}) - V_\upsilon(\phi_{\text{imit}})]$

How “rational” does ϕ_{demo} appears to be in pursuing (the “ideal”) υ?

- F = a bounded rational planner
- $\theta_{\text{norm}} = \upsilon$
- θ_{desc} = measures of “rationality”
- Appropriate inverse planner G

- Inverse rational bounded control

Bounded rational control

Uncertain knowledge of the environment

- **Unbiased prior:** $\tilde{\sigma} \in \Delta(\mathcal{T}, \mathcal{O})$
- **Biased specification policy:** $\sigma(z,u) \in \Delta(\mathcal{T}, \mathcal{O})^{\mathcal{Z},\mathcal{U}}$

Recognition policy is given in terms of specification policy

- $\varrho(z'|z,u) = \mathbb{E}_{\tau,\omega\sim\sigma(z,u),x'\sim\omega}[\rho_{\tau,\omega}(z'|z,u,x')]$
- $\rho_{\tau,\omega}$ **could be Bayesian inference under perfect knowledge** τ, ω

Bounded rational planner:

maximize $\mathbb{E}_{\pi,\rho,\sigma}[\sum_t v(s_t,u_t)]$ **s.t.**

$$\mathbb{E}_z[D_{\mathrm{KL}}(\pi(\cdot\,|z)||\tilde{\pi})] < A \longrightarrow \text{Decision complexity}$$

$$\mathbb{E}_{z,u}[D_{\mathrm{KL}}(\sigma(\cdot\,|z,u)||\tilde{\sigma})] < B \longrightarrow \text{Specification complexity}$$

$$\mathbb{E}_{z,u,\tau,\omega}\left[D_{\mathrm{KL}}\left(\varrho_{\tau,\omega}(\cdot\,|z,u)||\tilde{\varrho}\right)\right] < C \longrightarrow \text{Recognition complexity}$$

Bounded rational control

Value iteration:

$$V(z) \leftarrow \mathbb{E}\left[v(s,u) + \gamma V(z') \underbrace{- \alpha \log \frac{\pi(u|z)}{\tilde{\pi}(z)} - \beta \log \frac{\sigma(\tau,\omega|z,u)}{\tilde{\sigma}(z,u)} - \eta \log \frac{\varrho_{\tau,\omega}(z'|z,u)}{\tilde{\varrho}(z')}}_{\text{Complexity terms}}\right]$$

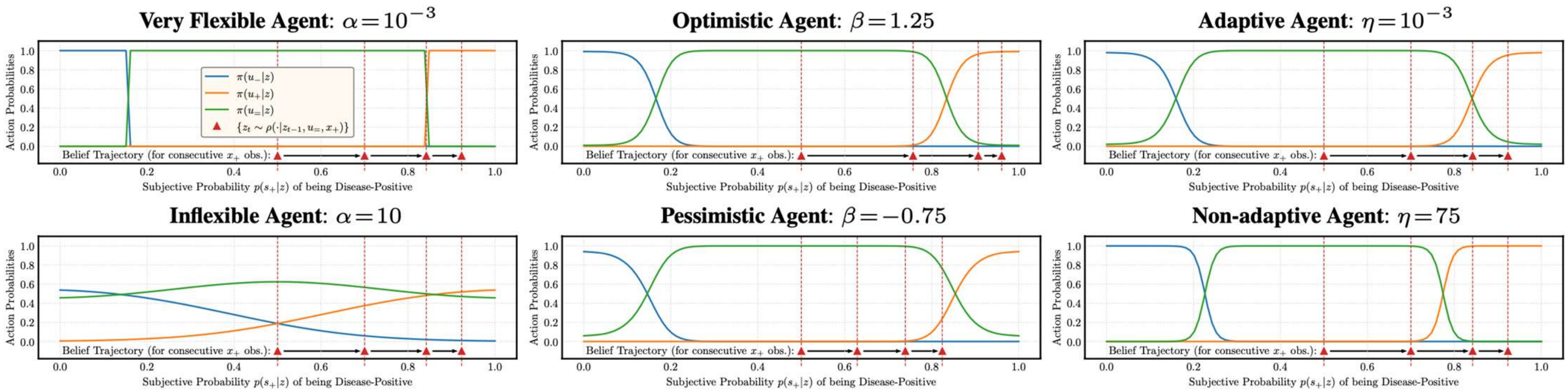
- $1/\alpha$ is a measure of flexibility
- $1/\beta$ is a measure of optimism
- $1/\eta$ is a measure of adaptivity

Flexibility, optimism, adaptivity

- **Observations:** negative (x_-), positive (x_+)
- **Actions:** monitor ($u_=$), negative diagnosis (u_-), positive diagnosis (u_+)
- **Utility:** 10 for correct diagnoses, -36 for incorrect diagnosis, -1 for monitoring

Flexibility, optimism, adaptivity

- Observations: negative (x_-), positive (x_+)
- Actions: monitor ($u_=$), negative diagnosis (u_-), positive diagnosis (u_+)
- Utility: 10 for correct diagnoses, -36 for incorrect diagnosis, -1 for monitoring



(a) Effect of Flexibility, for a neutral ($\beta = 10^3$), adaptive ($\eta = 10^{-3}$) agent

(b) Effect of Optimism, for a flexible ($\alpha = 0.5$), adaptive ($\eta = 10^{-3}$) agent

(c) Effect of Adaptivity, for a flexible ($\alpha = 0.5$), neutral ($\beta = 10^3$) agent

cf. behavioral inconsistency

cf. over-/underreaction

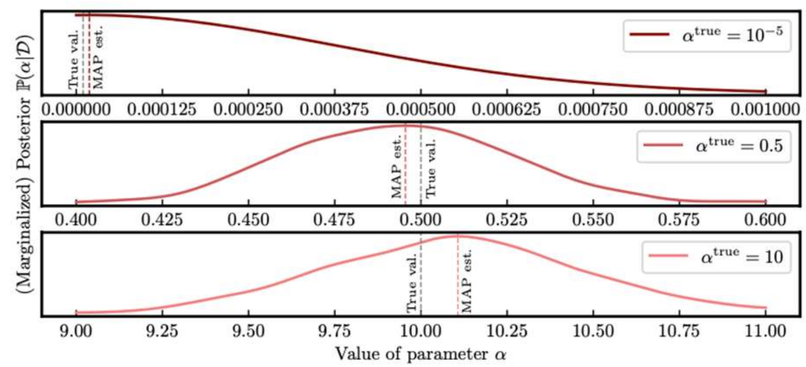
cf. base rate neglect/confirmation bias

Inverse bounded rational control

How "rational" does ϕ_{demo} appears to be in pursuing (the "ideal") υ?

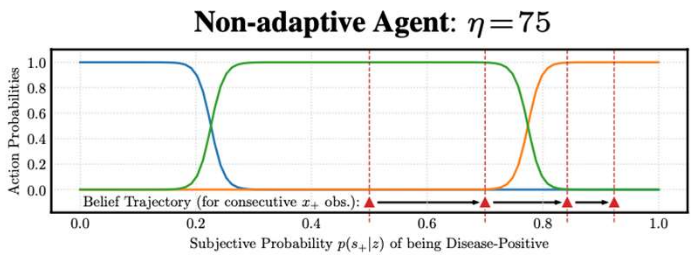
Inverse bounded rational control:

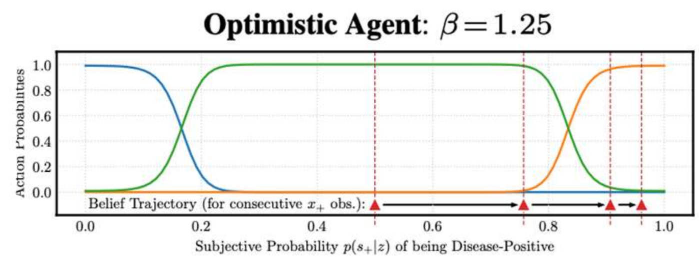
- $F =$ **the bounded rational planner**
- $\theta_{\text{norm}} = \upsilon$
- $\theta_{\text{desc}} = \alpha, \beta, \eta$
- $\mathcal{G}(\phi_{\text{demo}}, \phi_{\text{imit}}) = \mathbb{E}_{x,u \sim \phi_{\text{demo}}}\left[\mathbb{P}_{\phi_{\text{imit}}}(u_{1:T} || x_{1:T})\right]$



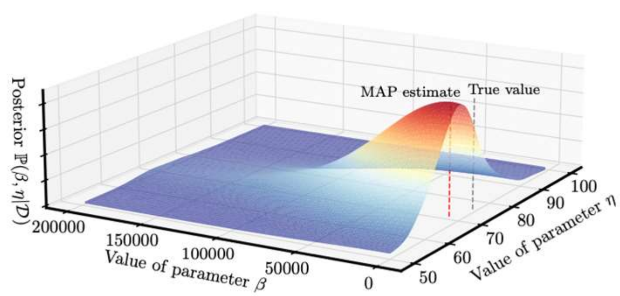
Learned α for various levels of flexibility

Differentiating non-adaptivity and optimism

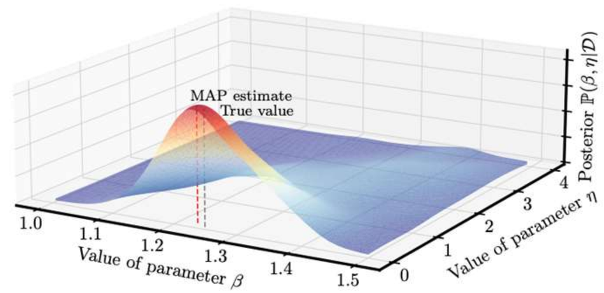




Non-adaptivity and optimism lead to similar behavior.



Learned β, η for Non-adaptive Behavior



Learned β, η for Optimistic Behavior

Utilities estimated by IRL:				
	10	for correct diagnoses	10	for correct diagnoses
	-26±3	for incorrect diagnoses	-27±3	for incorrect diagnoses

Illustrative use of IDM

IDM can be used as an investigative device for auditing and understanding human decision-making

Environment:

Diagnosing Alzheimer's disease
When to order an MRI?

- **MRIs are informative but costly**

$S = \{\text{NL}, \text{MCI}, \text{Dementia}\}$
$A = \{\text{MRI}, \text{No MRI}\}$
$Z = \text{Cognitive test results} \times \text{MRI outcomes}$

ADNI dataset

Pessimism when diagnosing Alzheimer's

Diagnosis of Alzheimer's:

When to order an MRI?

- **MRIs are informative but costly**

$\beta = 3.86$ **for all patients**

Clinicians appear to be significantly less optimistic when diagnosing:

- **patients with the ApoE4 genetic risk factor** $(\beta = 601.74)$
- **female patients** $(\beta = 920.70)$
- **patients aged >75** $(\beta = 2265.30)$

Our other work within the IDM framework

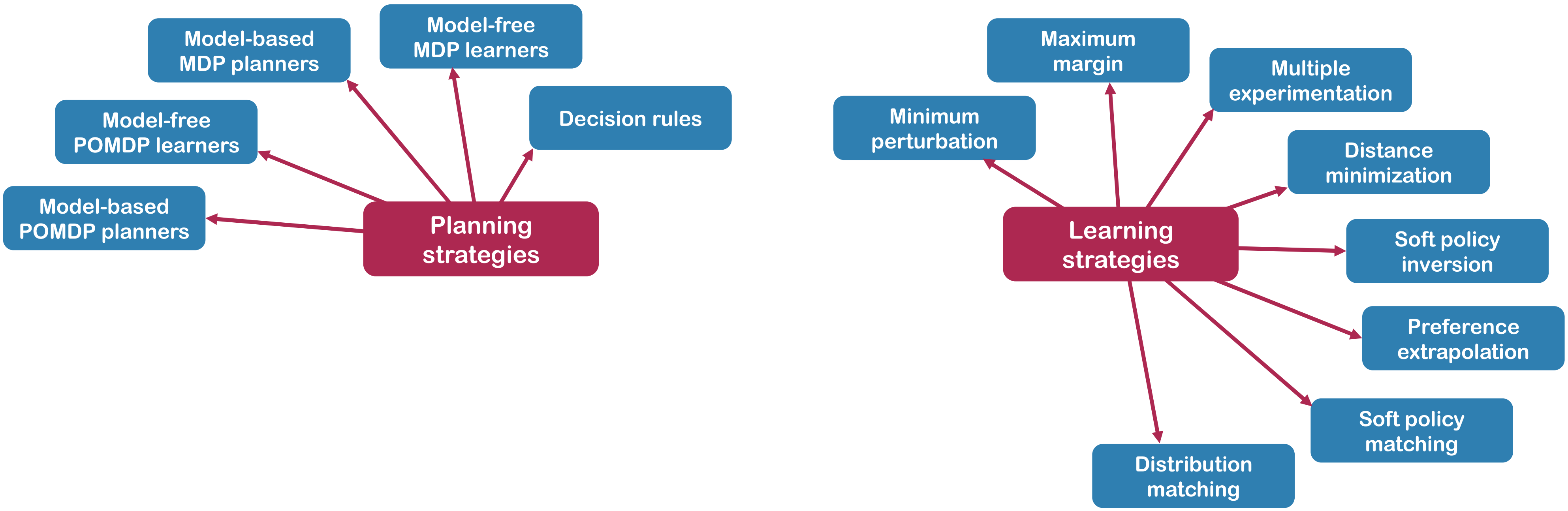
Method	Goal / motivating question	Planner	Normative params.	Descriptive params.
IAS (ICML'20)	How "timely" does the agent make decisions?	Timely active sensing	Deadline, cost of acquisition	Importance of accuracy, speed, efficiency
AVRIL (ICLR'21)	What reward function does the agent optimize?	RL planner	-	Reward function
CIRL (ICLR'21)	How important are various counterfactuals in making decisions?	Counterfactual RL planner	Counterfactuals	Importance weights
INTERPOLE (ICLR'21)	What are the subjective beliefs of the agent?	Policies based on decision boundaries	Interpretable state space	Decision dynamics & decision boundaries
IBRC (ICML'21)	How rational the agent behaves relative to an "ideal" reward function?	Bounded rational planner	"Ideal" reward function	Flexibility, optimism, adaptivity
ICB (submitted to NeurIPS'21)	How does behavior evolve over time?	Contextual bandit strategies	-	Time-varying beliefs over reward functions

IDM defines a broad class of potential studies in behavior representation learning

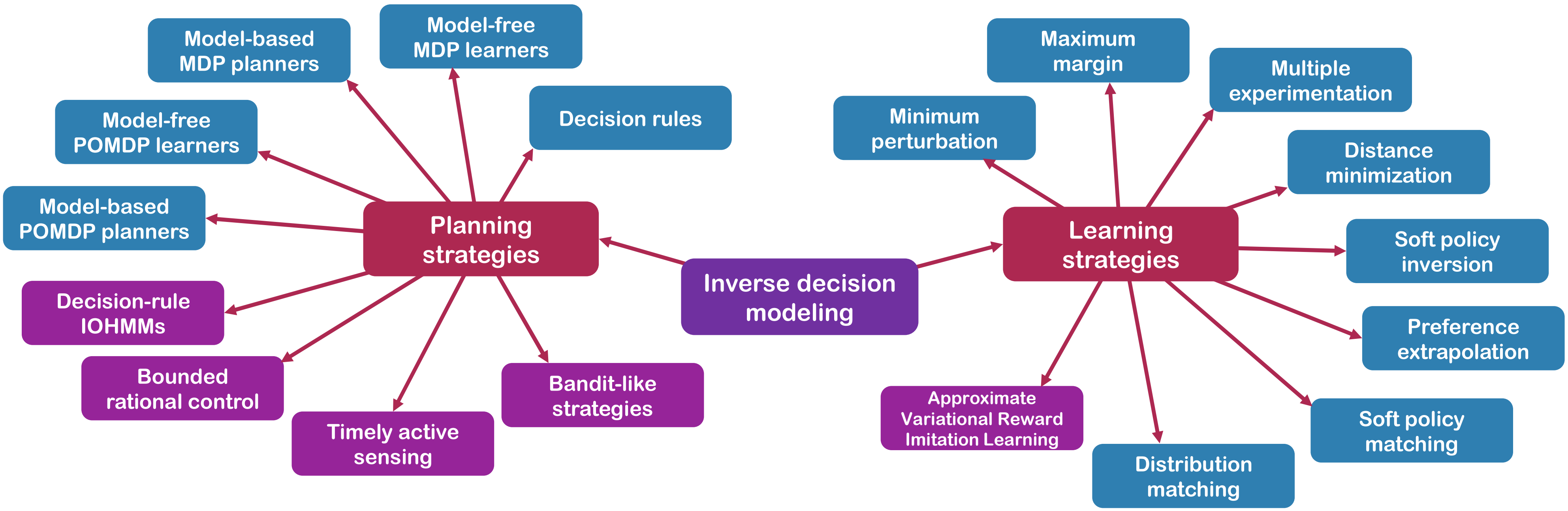
Previous works	Partially controllable	Partially observable	Purposeful behavior	Subjective dynamics	Action stochasticity
	Replacing & Outperforming humans				
Behavioral cloning	✓	✓	✗	✗	✓
Subjective behavioral cloning	✓	✓	✗	✓	✓
Deterministic distribution matching	✓	✗	✗	✗	✗
Stochastic distribution matching	✓	✗	✗	✗	✓
Deterministic IRL	✓	✗	✓	✗	✗
Stochastic IRL	✓	✗	✓	✗	✓
Subjective IRL	✓	✗	✓	✓	✓
Risk sensitive IRL	✓	✗	✓	✓	✗
Deterministic partially-observable IRL	✓	✓	✓	✗	✗
Stochastic partially-observable IRL	✓	✓	✓	✗	✓
Subjective partially-observable IRL	✓	✓	✓	✓	✓
Maximum entropy IRL	✓	✗	✓	✗	✓
Subjective maximum entropy IRL	✓	✗	✓	✓	✓

					Understanding humans			
Inverse decision model	**Partially controllable**	**Partially observable**	**Purposeful behavior**	**Subjective dynamics**	**Action stochasticity**	**Knowledge uncertainty**	**Decision complexity**	**Specification complexity**	**Recognition complexity**
Behavioral cloning	✓	✓	✗	✗	✓	✗	✗	✗	✗
Subjective behavioral cloning	✓	✓	✗	✓	✓	✗	✗	✗	✗
Deterministic distribution matching	✓	✗	✗	✗	✗	✗	✗	✗	✗
Stochastic distribution matching	✓	✗	✗	✗	✓	✗	✗	✗	✗
Deterministic IRL	✓	✗	✓	✗	✗	✗	✗	✗	✗
Stochastic IRL	✓	✗	✓	✗	✓	✗	✗	✗	✗
Subjective IRL	✓	✗	✓	✓	✓	✗	✗	✗	✗
Risk sensitive IRL	✓	✗	✓	✓	✗	✓	✗	✗	✗
Deterministic partially-observable IRL	✓	✓	✓	✗	✗	✗	✗	✗	✗
Stochastic partially-observable IRL	✓	✓	✓	✗	✓	✗	✗	✗	✗
Subjective partially-observable IRL	✓	✓	✓	✓	✓	✗	✗	✗	✗
Maximum entropy IRL	✓	✗	✓	✗	✓	✗	✓	✗	✗
Subjective maximum entropy IRL	✓	✗	✓	✓	✓	✗	✓	✗	✗
Inverse bounded rational control	✓	✓	✓	✓	✓	✓	✓	✓	✓

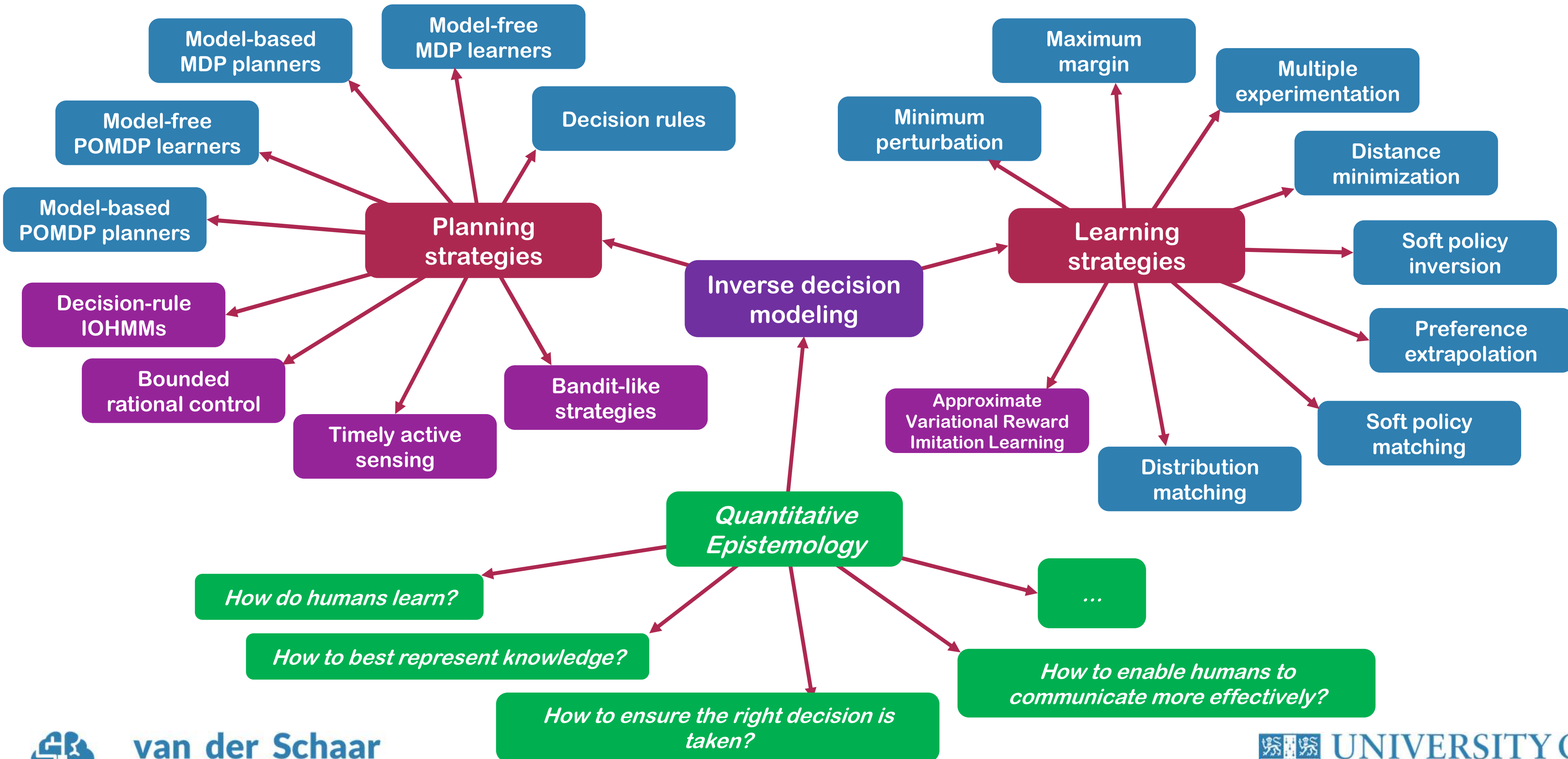
Quantitative Epistemology: New ML needed



Quantitative Epistemology: New ML needed



Quantitative Epistemology: New ML needed



van_der_Schaar \LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

Quantitative epistemology

- A new human-machine partnership
- A new field of multi-disciplinary research
- Partnering with humans to empower them, not to replace them!

The Standard ML Agenda

A standard ML scenario: no human agency

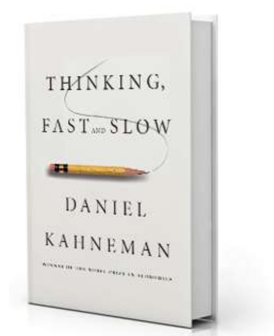
Screening and presentation	Diagnosis	Prognosis	Treatment	Follow-up

Predictions and recommendations

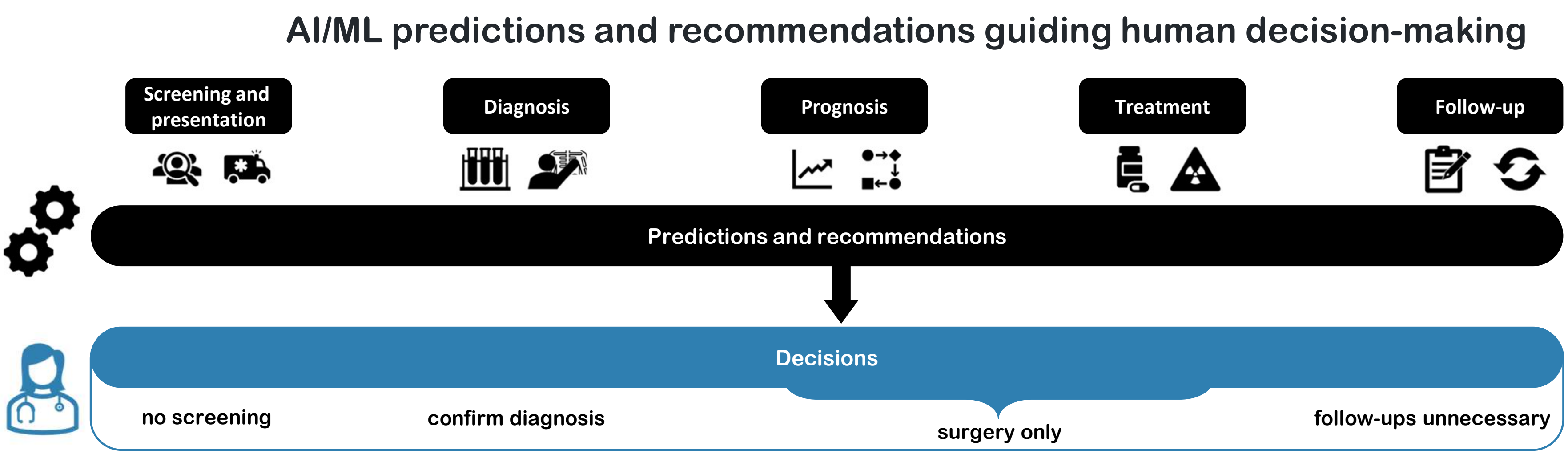
"should be screened early"	"stage 1 breast cancer, 1 cm, 2 positive nodes"	"10-year survival likelihood of 90%"	"low survival benefit from chemotherapy"	"mammograms every 6 months"

Decisions

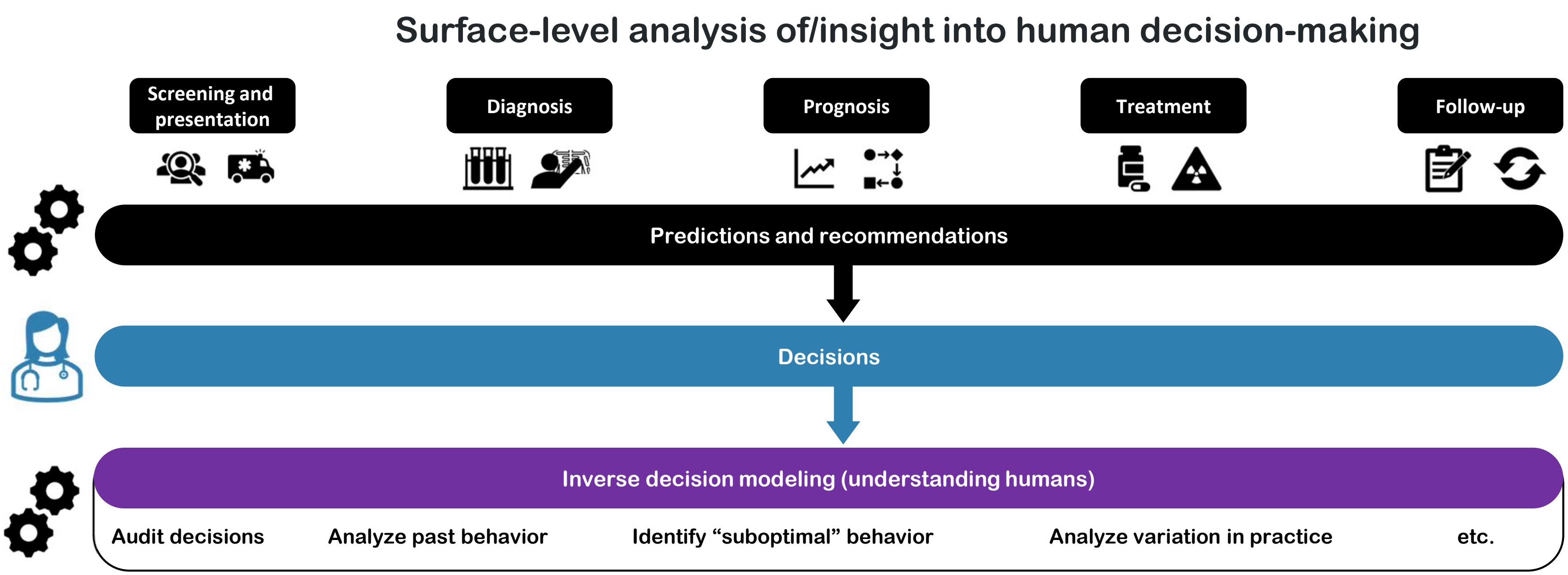
Learn how humans make decisions;
incorporate this into the design of more human-like AI/ML;
REPLACE & OUTPERFORM HUMANS



Standard Decision Support

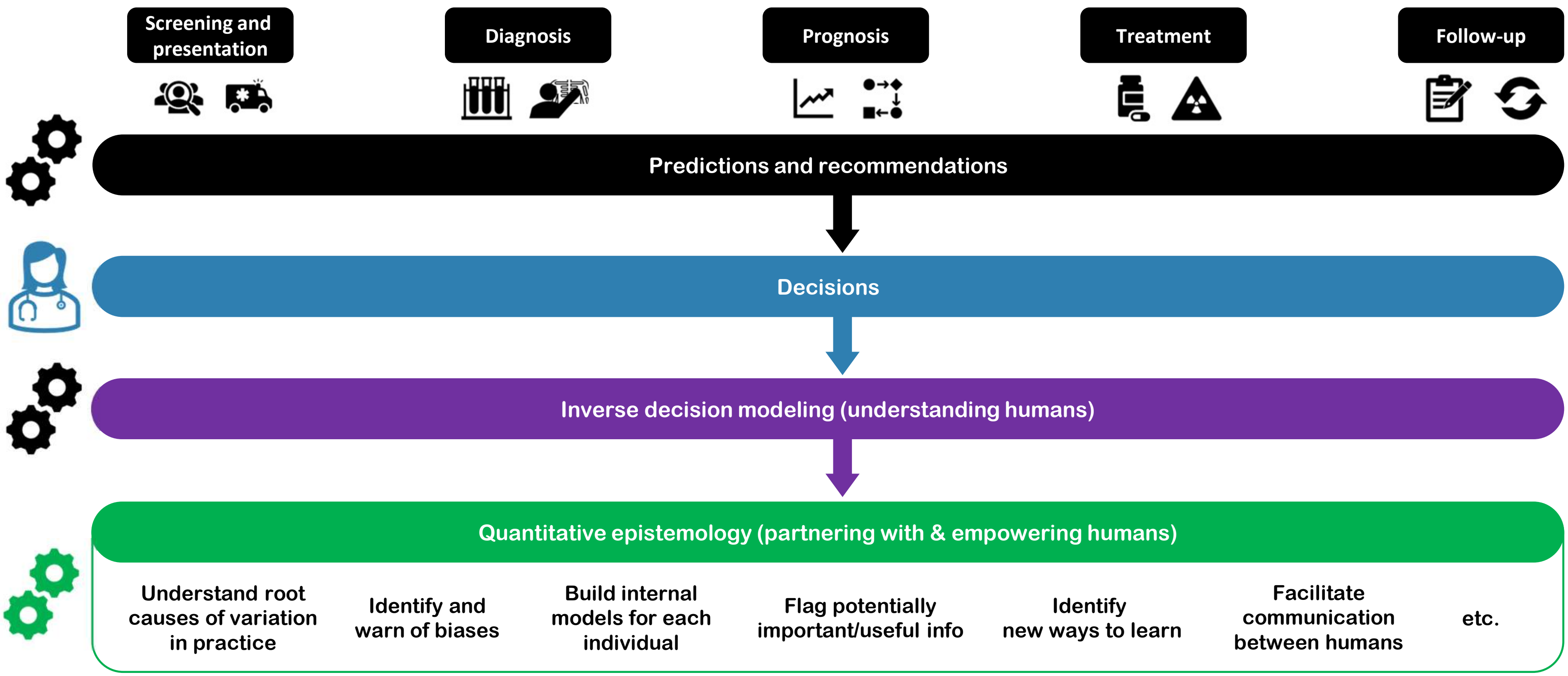


Inverse decision modeling (vdS-Lab)



Quantitative Epistemology (vdS-Lab)

Extracting actionable meaning from analysis of decision-making...



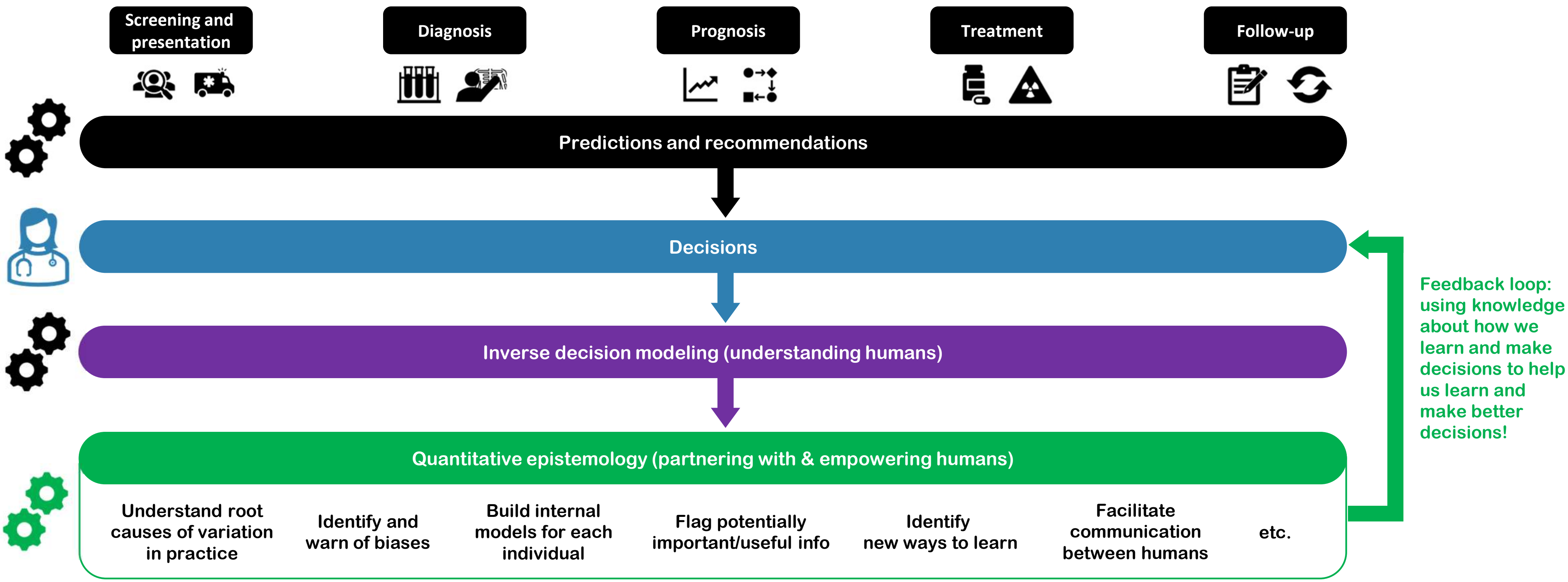
van_der_Schaar \LAB

vanderschaar-lab.com

UNIVERSITY OF CAMBRIDGE

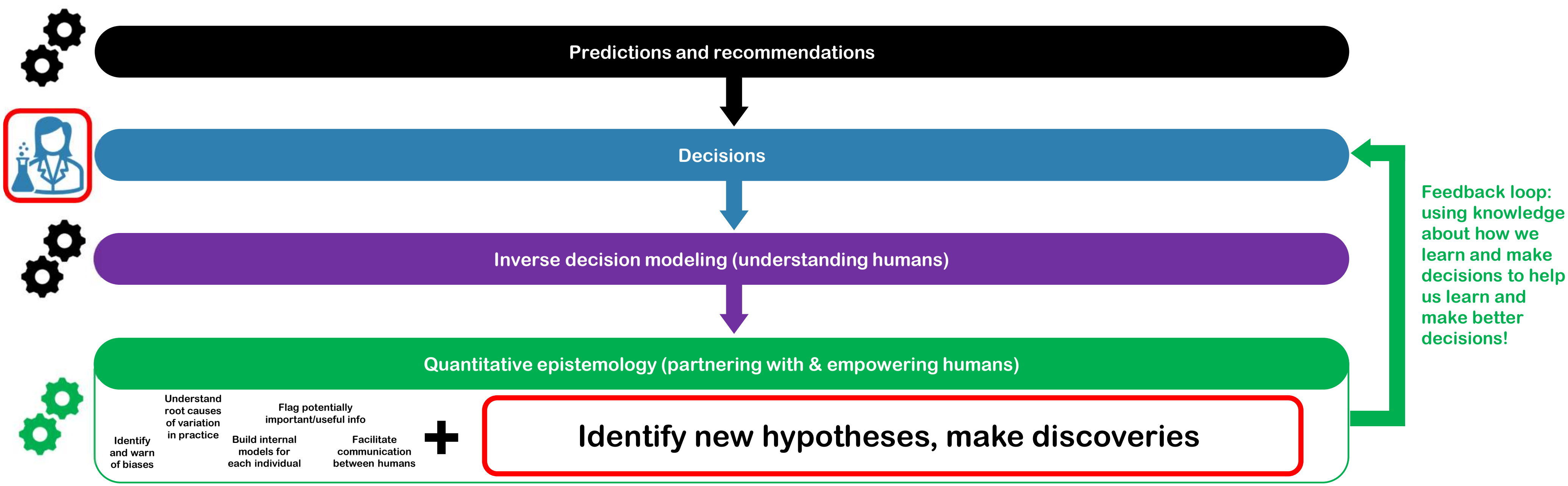
Quantitative Epistemology (vdS-Lab)

… creating an <u>empowering loop</u> that maximizes human agency and helps us make better decisions



Quantitative Epistemology (vdS-Lab)

For the researcher: new hypotheses and discoveries!



For more information & updates

vanderschaar-lab.com

→ Research pillars
→ Quantitative epistemology