

最終講義 (Final)

～「飴と鞭」の戦略：AI 育て・実践編～

■日時:

2026年2月28日(土) 14:30～16:30 最終講義
17:00～ 懇親会

■場所:

講義：工学・情報学系講義棟 1階 イノベーションシアター / 懇親会：松韻会館

*詳細・参加申込み

<https://sites.google.com/view/finallecturesarai2026/%E3%83%9B%E3%83%BC%E3%83%A0?authuser=0>

■対象:

どなたでも（学生、教職員、OB、一般の方、文系・理系を問わず）

■概要:

AI 設計の真髄は、高度な計算方法より、むしろ「何を良しとするか」という人間の知性の反映にあると思ってきた。すなわち、AI は人を映す鏡であり、人は AI を通して自らの価値観を省みることができる。しかし、人間の知性は、言葉にできない暗黙知として行動に潜んでおり、そのまま数式に書き下せる代物ではない。

私(荒井)は、成功時の「やった!」、失敗時の「しまった!」という試行錯誤の体験、そしてそれを導く「報酬(あめと鞭)」こそが知の源泉であると考え、強化学習を軸に、多様な価値や他者との調和を探る研究を進めてきた。本講義では、AI が人間の行動から目的や判断基準そのものを学ぶ「逆強化学習」の視点などを通じて、AI がいかに人の価値観を映し出すかを解説する。

最後に、報酬設計という難問と向き合う中で私が何を学び、次世代に何を託したいか、そして私が次にやりたいことを宣言し、有言実行のトリガーとする。

■講演者:

荒井 幸代 (工学研究院・教授)

▶ [研究室ウェブサイト](#) ▶ [researchmap](#)