

PRESS RELEASE (2025/12/17)

## リビングラボを用いて思考実験と経験的洞察の統合で新学際研究領域 「責任あるロボティクス」へ

～高次哲学的考察と HRI 実証研究を統合した革新的な AI 倫理影響評価方法論の提案～

### ポイント

- ① AI のペーシング問題(※1)によって引き起こされる智能ロボット規制のギャップを克服するための革新的なデザイン主導型ガバナンス(※2)として、ロボット倫理の標準化と規制におけるリビングラボ(※3)の新たな応用を探索することを目的としています。
- ② 本研究の内容は現実世界において発生するロボットの倫理的、法的、社会的課題 (ELSI) を対象として、九州大学における [LOVOT](#) 人間型ロボットのプライバシー保護問題と東北大学における「優しい嘘」をつく智能ロボットの倫理配慮問題に基づくリビングラボを用いて高次哲学的考察と HRI 実験(※4)を統合した業界初ハイブリッド AI 倫理影響評価方法論を提案しました。
- ③ 九州大学高等研究院と東北大学学際科学フロンティア研究所の戦略的連携で、福岡と仙台の両ハブの上に構築された総合知・超学際組織リサーチ・ネットワーク [ROBOLAW.ASIA](#) を通して IEEE P7017 リビングラボと青葉山リビングラボを活用し新学際研究領域「責任あるロボティクス」の促進が期待されます。

### 概要

ロボットが日常生活に統合されると、人間との相互作用に起因する複雑な倫理的、法的、社会的課題 (ELSI) が生じます。ソーシャルロボットは、文化的規範、感情、社会的手がかりが豊富な環境で動作するため、プライバシー、信頼、安全性に関する重要な問題を提起します。本研究では、九州大学高等研究院の翁岳暄准教授(東北大学学際科学フロンティア研究所(クロスアポイントメント))、同大学大学院法学府の David Torabi (デイビッド トラビ) 院生、オスロ大学大学院情報科学研究科の Jim Torresen (ジム トレセン) 教授、東北大学大学院工学研究科の董宗昊特任助教、東北大学大学院工学研究科の平田泰久教授らの協力のもとロボット倫理学という学際的な分野が、思考実験と実証研究を組み合わせたハイブリッドな方法論的概念を通じて、これらの課題をどのように対処できるかを探ります。思考実験は、倫理的なジレンマを体系的に分析するための枠組みを提供します。一方、実証的な方法は、こうした理論を現実世界で検証し、より洗練させるための洞察を与えます。

本論文では特に、ロボットが倫理的な期待や法的基準に適合することを確保するために、倫理設計の原則をロボット設計に組み込み、テストする場として「リビングラボ」を活用する重要性を強調しています。

本研究成果はアメリカの雑誌「IEEE Robotics & Automation Magazine」に 2025 年 12 月 11 日付で掲載されました。

### 研究者からひとこと：

責任あるロボティクスには学際的協働が不可欠であり、技術革新だけでは社会的課題に対応できないことが再確認されました。産業ロボット中心から生活環境に組み込まれる社会的ロボットへの研究転換が進む中、デザイン主導型ガバナンスの重要性が増えるはずです。 翁 岳暄 (ウエン ユエ シュアン)

## 【研究の背景と経緯】

AI 倫理基準とは、簡単に言えば、更新が容易で柔軟性があり、人々の行動を規制する強制力を持たないソフトローの一種です。法律の制定と改正にかかる時間は AI 技術の更新ペースよりもはるかに遅いため、AI 規制において法律に代わるソフトローの活用は近年ますます注目を集めています。本研究では、ソフトローの中でも IEEE 7000 シリーズの AI 倫理基準に焦点を当て、その理論と知能ロボットに関する ELSI リスク管理への応用を研究してきました。

## 【研究の内容と成果】

思考実験とは、制御された精神的枠組みの中で特定の変数を分離することにより、原則や意思決定の含意を探るために用いられる仮説的なシナリオです。哲学で伝統的に用いられてきたように、思考実験は研究者が道徳的な直観を探り、倫理理論を検証し、現実世界での試さずとも概念の境界を明確にすることを可能にします。ロボット倫理の文脈において、思考実験は、自律システムが遭遇する可能性のある複雑なジレンマを想定し、そのときどのような規範に従うべきかを分析するのに役立ちます。こうした方法の価値は、特に現実世界にロボットを導入するのがまだ早い場合や、リスクを伴う可能性がある状況で、事前に問題を予測し、深く考察できる点にあります。

本研究グループでは、ロボット倫理学という学際的な分野が、思考実験と実証研究を組み合わせたハイブリッドな方法論的概念を通じて、ロボット倫理の標準化におけるリビングラボの新たな応用を探求することを目的としています。具体的には、用語の定義、ケーススタディ、HRI 実験などが含まれます。中でも最も重要なのは HRI 実験であり、倫理的問題の定義、倫理的シナリオのシミュレーション、社会規範の検証などが含まれます。倫理的ロボットデザインのアプローチを探求し、その可能性と限界を検証し、リーガル・ヒューマン・ロボット・インタラクション (L-HRI) (※5)の実施におけるコンプライアンス・バイ・デザイン(※6)の有用性を実証します。研究の内容は現実世界において発生するロボット倫理的、法的、社会的課題 (ELSI) 問題を対象として、九州大学における LOVOT 人間型ロボットのプライバシー保護問題と東北大学における「優しい嘘」をつく知能ロボットの倫理配慮問題に基づくリビングラボを用いて高次哲学的考察と HRI 実験を統合した業界初ハイブリッド AI 倫理影響評価方法論を提案しました。

## 【今後の展開】

本研究では特に、ロボットが倫理的期待と法的基準に適合することを保証するため、倫理的設計原則をロボット設計にテストし、統合するための動的な環境としてのリビングラボの活用を強調しています。

【参考図】

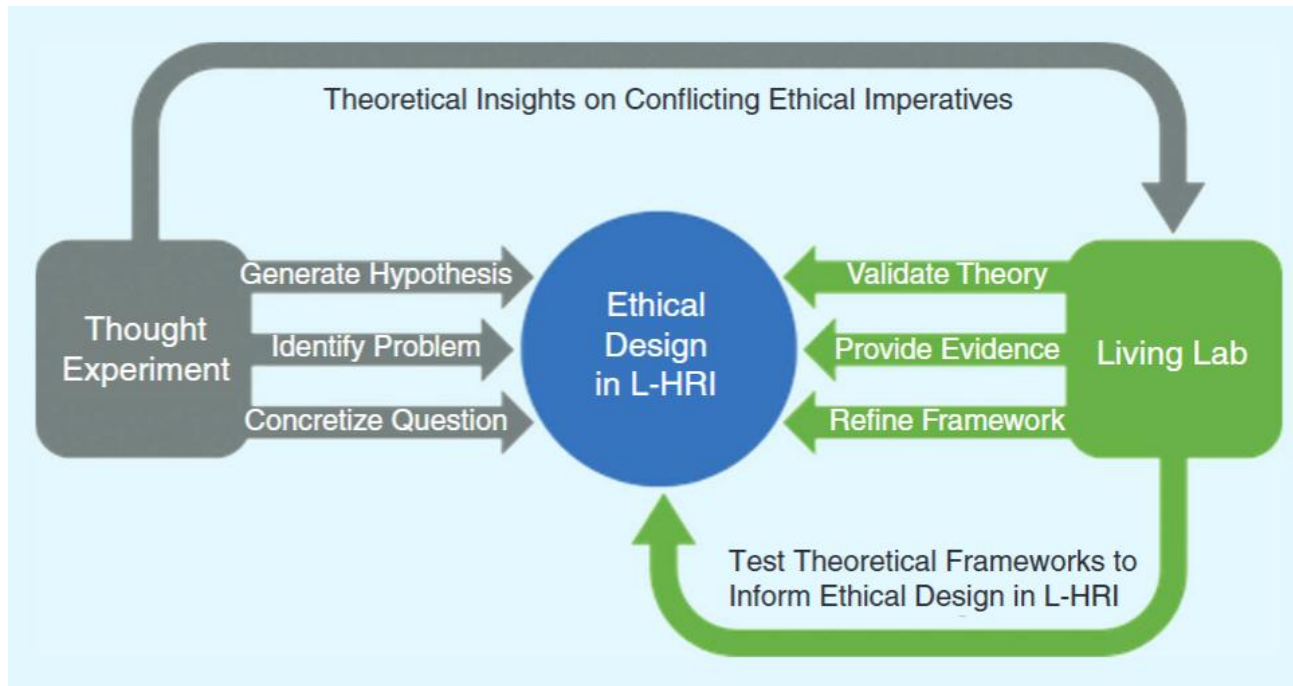


図 リビングラボにおける思考実験と実証研究を組み合わせた方法論的概念

【用語解説】

(※1) AI のペーシング問題

技術開発のスピードと、AI 対応技術の規制に対する立法措置の遅さとのミスマッチが明らかとなっている問題のこと。

(※2) デザイン主導型ガバナンス

AI 搭載型ソーシャルロボットがもたらす設計・倫理・ガバナンスの課題を多角的に検討。ソーシャルロボットを技術・行動・規範が相互作用する「社会技術システム」として捉え、開発初期段階から統合的視点を取り入れる必要性を重視する基本的な姿勢を強調。

(※3) リビングラボ

実際の生活環境の中で新しい技術やサービスを開発・検証するための手法。ユーザーや市民が開発プロセスに積極的に参加し、実環境でのフィードバックを提供。リビングラボが AI 規制において重要な理由は AI システムの実際の社会的影響を、管理された実環境で観察・評価でき、過度な規制によるイノベーション阻害を避けつつ、適切な安全措置を見出すための実験的アプローチを提供する。

(※4) HRI 実験（Human-Robot Interaction：ヒューマン・ロボット・インタラクション実験）

人間とロボット間の相互作用を研究・評価するための実験。HRI 実験は、より自然で効果的な人間とロボットの関係構築に向けた知見を提供し、将来のロボット技術の設計や実装に重要な役割を果たしている。

(※5) リーガル・ヒューマン・ロボット・インタラクション（L-HRI）

ロボットの設計、開発、使用のライフサイクルに、法的・倫理的基準の遵守を直接組み込む設計中心のアプローチ。

#### (※6) コンプライアンス・バイ・デザイン

法的・倫理的基準をロボットの設計、機能、行動に直接組み込むアプローチ。その目的は、機械が人間のようにハイレベルな法的・倫理的基準を直接に認識処理せずとも、ロボットが最初から法的要件や倫理基準に準拠するように設計されていることを保証すること。これは、ロボットの行動を誘導するセーフガード、行動制約、技術的手段を組み込むことによって可能になり、許容される法的・倫理的境界の範囲内でロボットが動作することを保証する。

#### 【謝辞】

本研究は JST 創発的研究支援事業（JPMJFR222C）、JST ムーンショット型研究開発事業（JPMJMS2034）の助成を受けたものです。

#### 【論文情報】

掲載誌：IEEE Robotics & Automation Magazine

タイトル：Bridging Ethics and Reality: Integrating Thought Experiments and Empirical Insights in Robot Ethics

著者名：Y. -H. Weng, D. Torabi, J. Torresen, Z. Dong and Y. Hirata

D O I : 10.1109/MRA.2025.3584352

#### 【お問合せ先】

<研究に関すること>

九州大学 高等研究院 准教授 翁 岳暄（ウェン ユエ シュアン）

TEL：092-802-5432

Mail：weng.yueh.hsuan.647@m.kyushu-u.ac.jp

<報道に関すること>

九州大学 広報課

TEL：092-802-2130 FAX：092-802-2139

Mail：koho@jimu.kyushu-u.ac.jp