

AI 技術に対する社会の態度を「オクタゴン測定」で数値化

2022 年 1 月 10 日

東京大学国際高等研究所カブリ数物連携宇宙研究機構 (Kavli IPMU)

金沢大学人間社会研究域

1. 発表概要

東京大学国際高等研究所カブリ数物連携宇宙研究機構 (Kavli IPMU) の横山広美 (よこやま ひろみ) 教授を中心とする東京大学、金沢大学のメンバーからなる研究グループは、先端技術に対して人々が感じる倫理的・法的・社会的課題 (Ethical, Legal and Social Issues: ELSI) のレベルを可視化するプロジェクトを行っています。その最初の試みとして、先行研究で抽出された AI (人工知能) 倫理に共通する8つの観点に基づき、それを数値化する「オクタゴン測定」という新たな手法を提案しました。また、オクタゴン測定を用いて、異なる状況の4つの AI 技術のシナリオの倫理問題のレベルを測る研究を行いました。AI の倫理については各国がガイドラインを作成していますが、人々が感じる AI 技術の倫理的問題レベルを数値化する研究は世界でもほとんど行われておらず、ユニークな試みと言えます。また、今回4つの AI 技術のシナリオを対象として分析を行なったところ、若者より年配者、男性よりも女性、AI の知識が少ない人よりも多い人の方が AI 技術に懸念を持つ傾向を明らかにしました。オクタゴン測定法は、異なる状況や国際間での比較など、AI 技術の倫理問題を可視化する手法として、広く使用するのに適しています。プロジェクトメンバーはこの測定が、広く AI 技術の開発時に利用できる手法になると期待しています。

本成果は、ヒューマンコンピュータインタラクションの国際的学術誌である *International Journal of Human-Computer Interaction* のオンライン版に 2022 年 1 月 10 日付 (日本時間 1 月 11 日午前 0 時) で公開されました。

2. 発表内容

AI 技術の利用がますます活発になる現在、各国はそれぞれにガイドライン作成を行っています。昨今では各ガイドラインの共通部分を見出す研究が行われています。特にハーバード大学の Jessica Fjeld 氏らによる 36 のガイドラインを分析した研究は、「個人のプライバシー」「説明責任」「安全性とセキュリティ (第三者からの侵害)」「透明性と説明可能性」「公平性と無差別」「人間による制御」「専門家の責任」「人間の価値の促進」という 8 つの共通のテーマを扱っていることを明らかにしています (注)。

研究グループでは、異なる状況の AI 技術における倫理的な問題を可視化し、開発段階から科学技術者がこれを十分に把握することが重要だと考えました。同時に、国際的に比較をして共通部分を見つけ、違いを認識し国際的な議論を促すことも重要です。研究グループは、Fjeld 氏ら

の8つのテーマは AI 倫理に関する問題をよく分類しているため、これを数値化できる軸として設定し、オクタゴン（8 角）軸と呼ぶことにしました。さらに、メリットとデメリットの両方を含む倫理的ジレンマを含む4つのシナリオ（シナリオ1: AI 歌手、シナリオ2: AI ショッピング、シナリオ3: AI 兵器、シナリオ4: AI 犯罪予防）を用意し、これらの倫理問題レベルを、オクタゴン軸を用いて測定しました。シナリオでは、「私は情報科学の研究者です」から始まり、AI を利用することによる利点と不安な点を述べ、「私はこの AI 研究を続けるべきでしょうか？」と回答者に問うています。

データはウェブ調査会社を通じて、2020 年 9 月 10 日から 9 月 14 日の期間で、20 代から 60 代まで人口構成比に従って日本国内の男女 1029 人（女性 510 人、男性 519 人）から取得しました。

その結果、特に AI 兵器のシナリオでは、「個人のプライバシー」を除く7つの観点すべてで倫理問題レベルが高いという回答が得られました（図 1）。さらに回答者の年齢や性別といったプロフィールを用いて解析をした結果、シナリオに関わらず年齢が高いほど否定的という結果が得られました。また、一部のシナリオや一部の観点では女性ほど否定的、知識が高い人ほど否定的といった結果が得られました。こうした傾向は、他の科学技術のリスクにおいても見られる傾向であり、AI 技術も同様の傾向があることを確認しました。今後、同じシナリオについてオクタゴン測定法を用いて異なる国で測定することにより、倫理レベルの国際比較を行うことができると考えられます。また、研究グループでは、より簡便な軸の開発も現在検討しています。

3. 発表雑誌

雑誌名: International Journal of Human-Computer Interaction

論文タイトル: Octagon measurement: public attitudes toward AI ethics

著者: Yuko Ikkatai (1), Tilman Hartwig (2), Naohiro Takanashi (3), Hiromi M Yokoyama (4)

著者所属:

- 1) Faculty of Human Sciences, Institute of Human and Social Sciences, Kanazawa University
- 2) Institute for Physics of Intelligence, The University of Tokyo, Tokyo, Japan
- 3) Executive Management Program Office, The University of Tokyo
- 4) Kavli Institute for the Physics and Mathematics of the Universe (Kavli-IPMU), The University of Tokyo, Kashiwa, Japan

4. 用語解説

注)

ハーバード大学の Jessica Fjeld 氏らが 2020 年に発表した下記論文による。

Fjeld et al. (2020). Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-Based Approaches to Principles for AI, Berkman Klein Center Research Publication No. 2020-1, 39

5. 問い合わせ先

(研究内容について)

横山 広美(よこやま ひろみ)

東京大学国際高等研究所カブリ数物連携宇宙研究機構 教授

E-mail: hiromi.yokoyama@ipmu.jp

※取材はメールおよび、1月7日(金) 10時-13時までの間、以下のZoomで対応可能です。

<https://ipmu.zoom.us/j/7758735226?pwd=L3FxTUxSZmt3YWxhSHIDNTQ2MlVjUT09>

プロジェクトウェブサイト: <http://member.ipmu.jp/hiromi.yokoyama/secom.html>

(報道対応)

東京大学国際高等研究所カブリ数物連携宇宙研究機構 広報担当 小森真里奈

E-mail: press@ipmu.jp TEL: 080-4056-2930 / 04-7136-5977

国立大学法人金沢大学 人間社会系事務部総務課総務係 山本

E-mail: n-somu@adm.kanazawa-u.ac.jp TEL: 076-264-5598

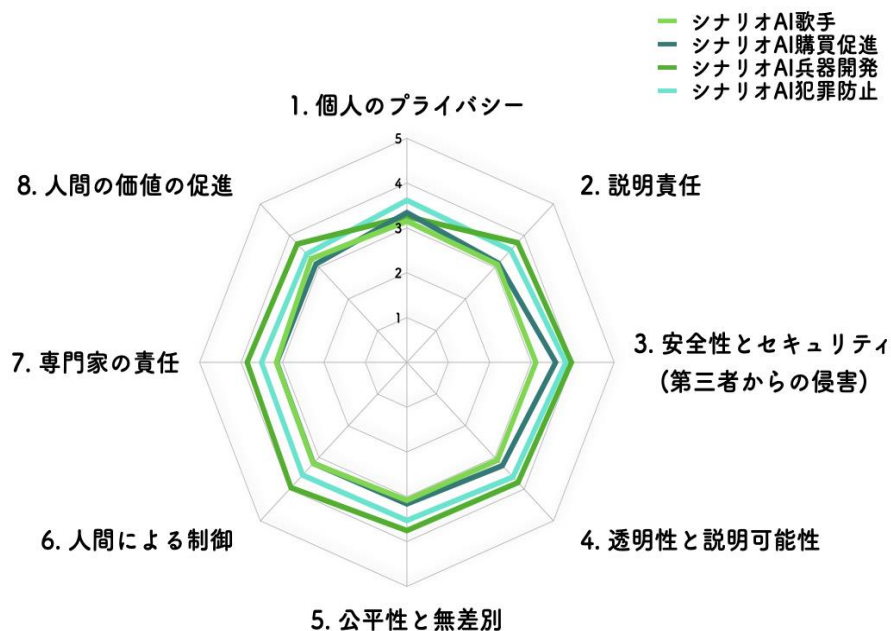


図1. 4つのシナリオの倫理問題レベルを測定するオクタゴン軸。日本で約 1029 人のデータを取

得した。回答者は8つの観点について5件法（現状でよい＝1、今よりやや配慮が必要である＝2、今よりかなり配慮が必要である＝3、今より非常に配慮が必要である＝4、今より極めて配慮が必要である＝5）で回答した。数値が大きいほど、倫理問題レベルが高い（問題である）ことを意味している。AI 兵器開発に対しては、「1. 個人のプライバシー」を除く7つの観点全てで倫理問題レベルが高いという回答が得られた。（Credit: Score ELSI プロジェクト）

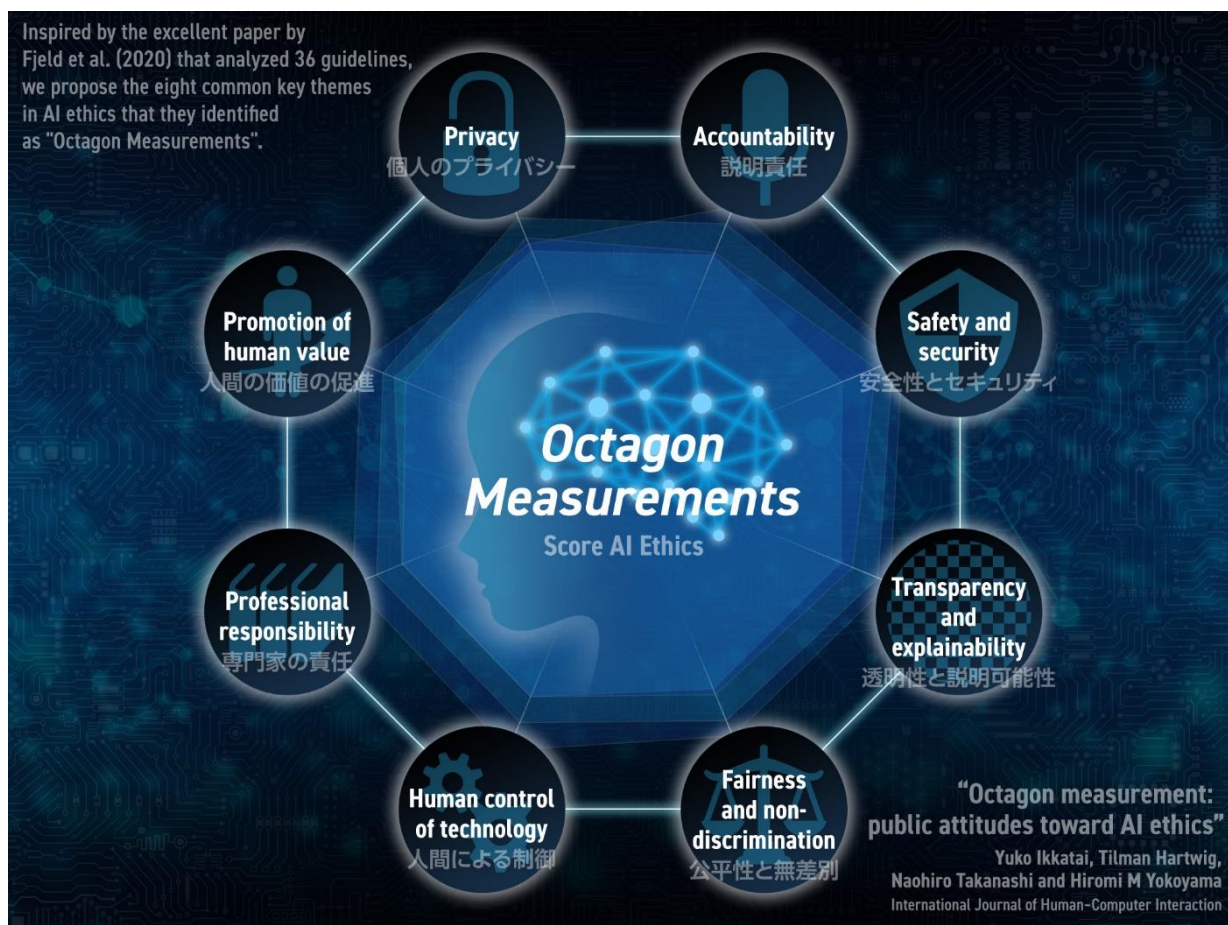


図2. 研究グループは、「オクタゴン測定」は AI 倫理を測定するのに適していると考えている。
(Credit: Score ELSI プロジェクト)