

科学技術社会論から見る AI倫理の現在



横山広美(科学技術社会論)

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・副機構長

<https://member.ipmu.jp/hiromi.yokoyama/yokoyama.html>



東京大学 国際高等研究所 カブリ数物連携宇宙研究機構



高エネルギー素粒子実験(ニュートリノ物理)、博士(理学)出身の
人文社会科学の研究者

Kavli IPMU: Center for data driven discovery



開発時に未来を考える義務

ELSI/RRI

Ethical, Legal and Social Issues

Responsible Research and Innovation

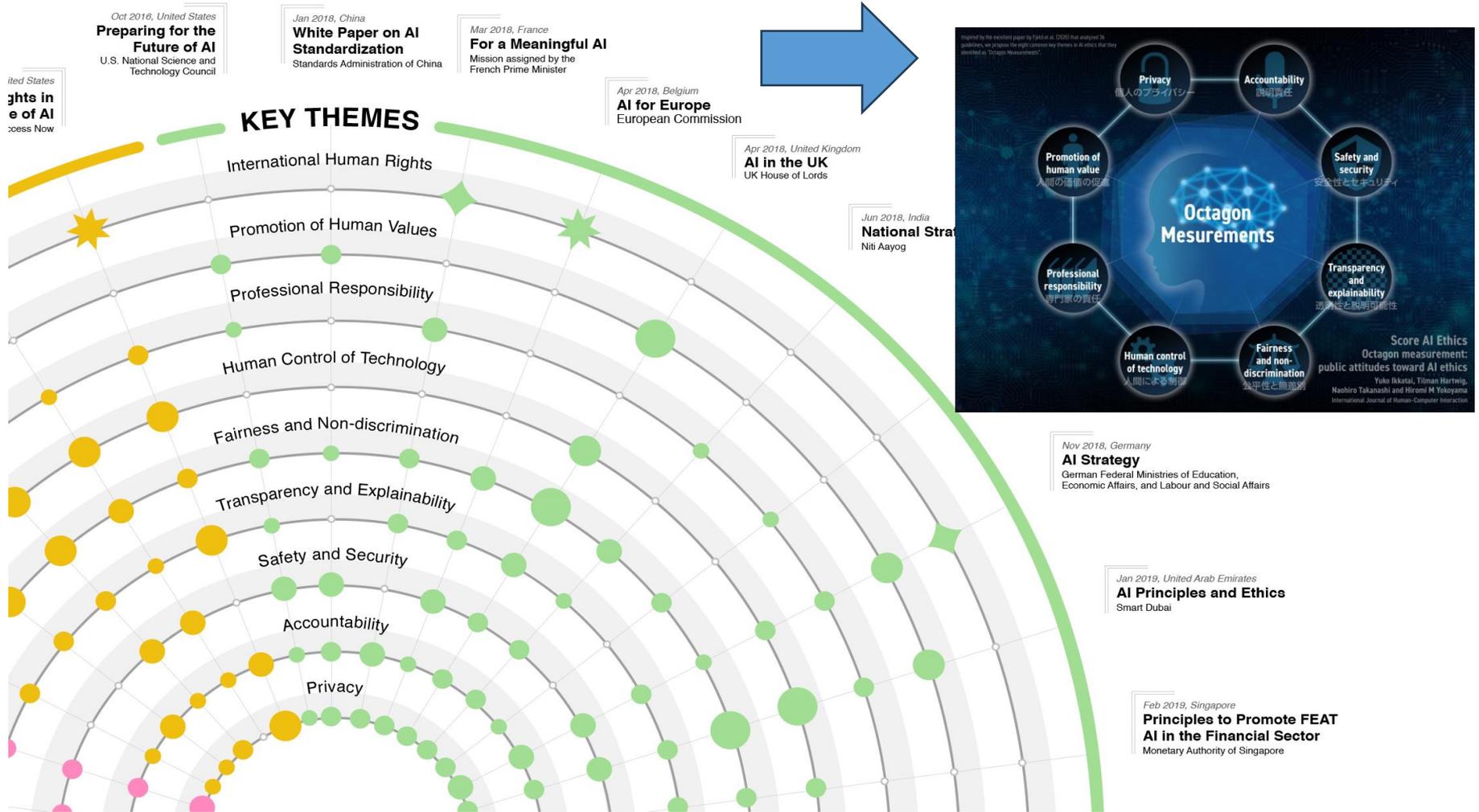


ドローン兵器の画像



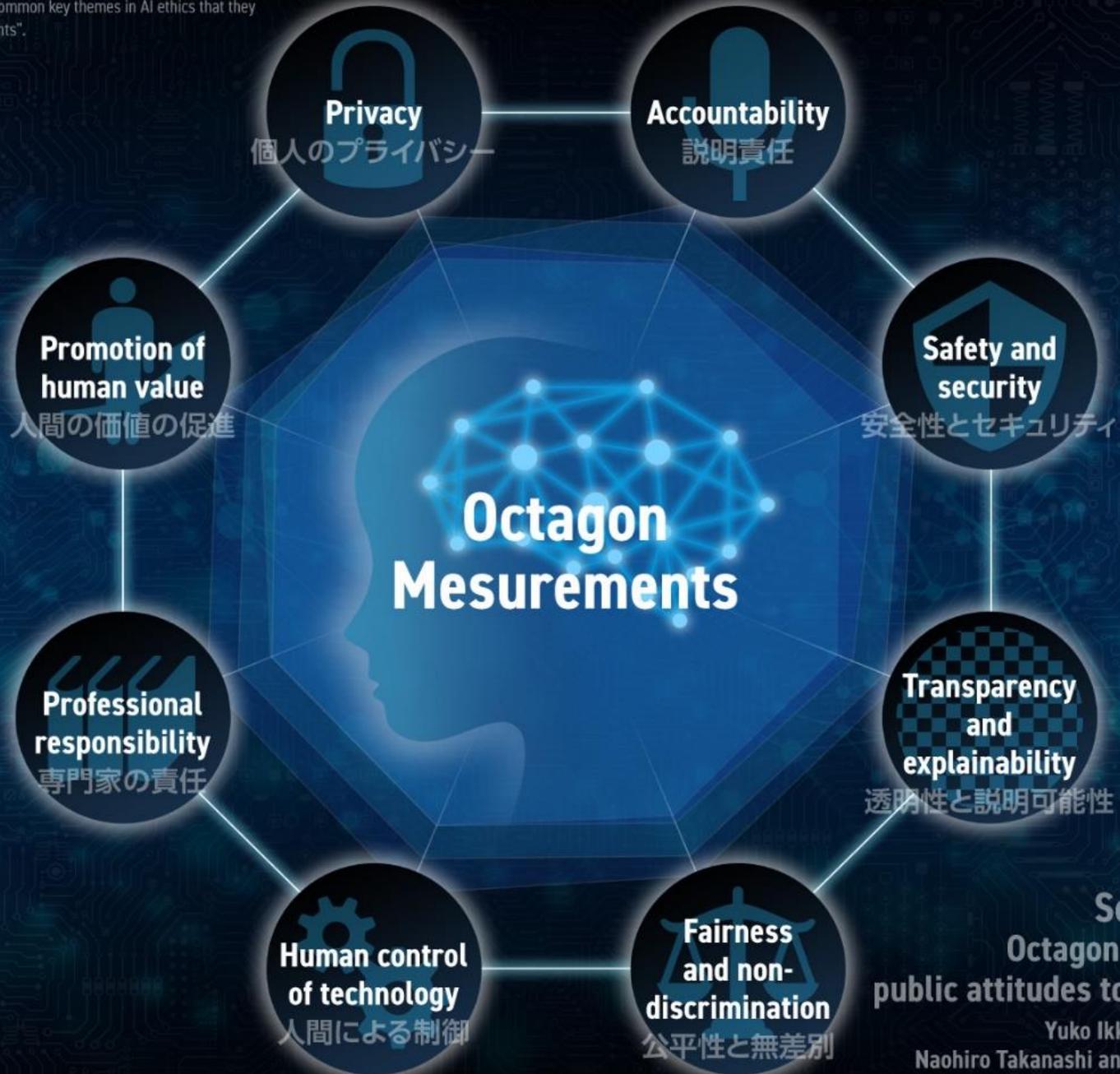
映画
オッペンハイマーの
ポスター

「オクタゴンメジャメント」を提案



世界中のガイドラインを分析したハーバードの分析を元に、世界共通の8つの軸を測定軸に用いる「オクタゴンメジャメント」を提案

Inspired by the excellent paper by Fjeld et al. (2020) that analyzed 36 guidelines, we propose the eight common key themes in AI ethics that they identified as "Octagon Measurements".



Score AI Ethics
Octagon measurement:
public attitudes toward AI ethics

Yuko Ikkatai, Tilman Hartwig,
Naohiro Takanashi and Hiromi M Yokoyama
International Journal of Human-Computer Interaction

「ELSIセグメント」を提案

物理学の手法を用いた解析で意見分布をグループ化
4つの場合がもっともポテンシャルが低くなる

Plot all 13,024 points of the four scenarios

K-means clustering shows that the four groups are ideal

The exact definitions of the 4 groups are as follows:

If (Ethics + Tradition + Law < 10):

Group 1

Elseif (Ethics + Tradition + Law > 16):

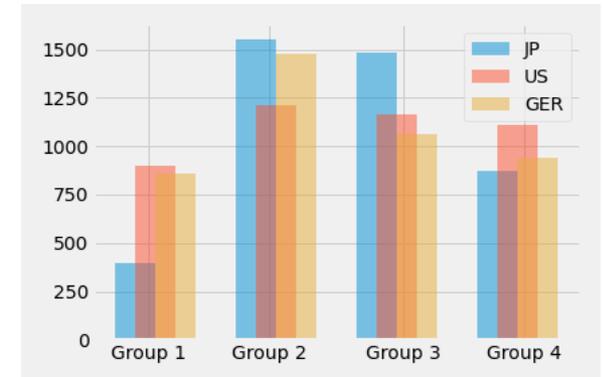
Group 4

Elseif (Law > 4):

Group 3

Else:

Group 2



ELSIを量的に測定し、意見の違いを可視化できる
「ELSIセグメント」を提案

12項目のELSIシートの中から抽出された、もっとも重要な3つの項目について、3軸として日米独の意見分布をプロット

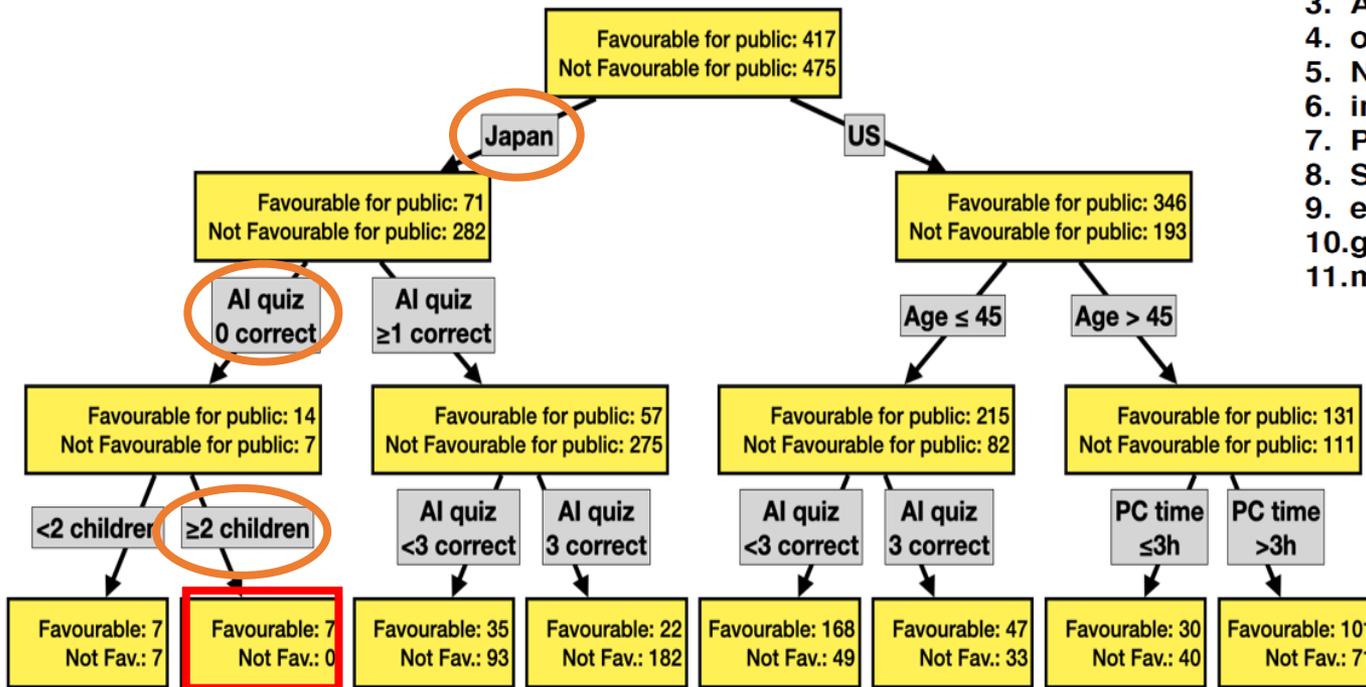
- 左下の青のグループがAI使用にポジティブ
- 右上のオレンジのグループがAI使用にネガティブ
- 中央の群青は法的対応を問題視、紫は問題視しない

Ikkatai, Y., Hartwig, T., Takanashi, & Yokoyama, HM. (2022) Segmentation of ethics, legal, and social issues (ELSI) related to AI in Japan, the United States, and Germany. *AI Ethics*

<https://doi.org/10.1007/s43681-022-00207-y>

AI(決定木解析)を用いて、 AI倫理の量的研究を実施

Scenario (c): Autonomous Weapons



1. country: 0.307
2. age: 0.295
3. AIquizCorrect: 0.180
4. occupation: 0.065
5. Nchildren: 0.049
6. income: 0.043
7. Pctime: 0.021
8. ScienceInterestGroup: 0.018
9. education: 0.011
10. gender: 0.007
11. married?: 0.003

日米のAI使用に対する
人々の態度を分析

←↑AI兵器：賛否を分ける重要ファクターは 1. 国、
2. 年齢（高いほど反対）、
3. AIについての知識レベル（高いほど反対）

科学技術社会論から見たChatGPTの特徴

科学論の観点から、他の科学技術と違う点

- ① 多点・多目的開発
- ② 開発、社会実装のスピードが非常に早い
- ③ 世界同時経験

→ 事前にリスクを検討対応する「**予防原則**」が困難
科学論からみると非常に危うい

ELSIと同走必要性

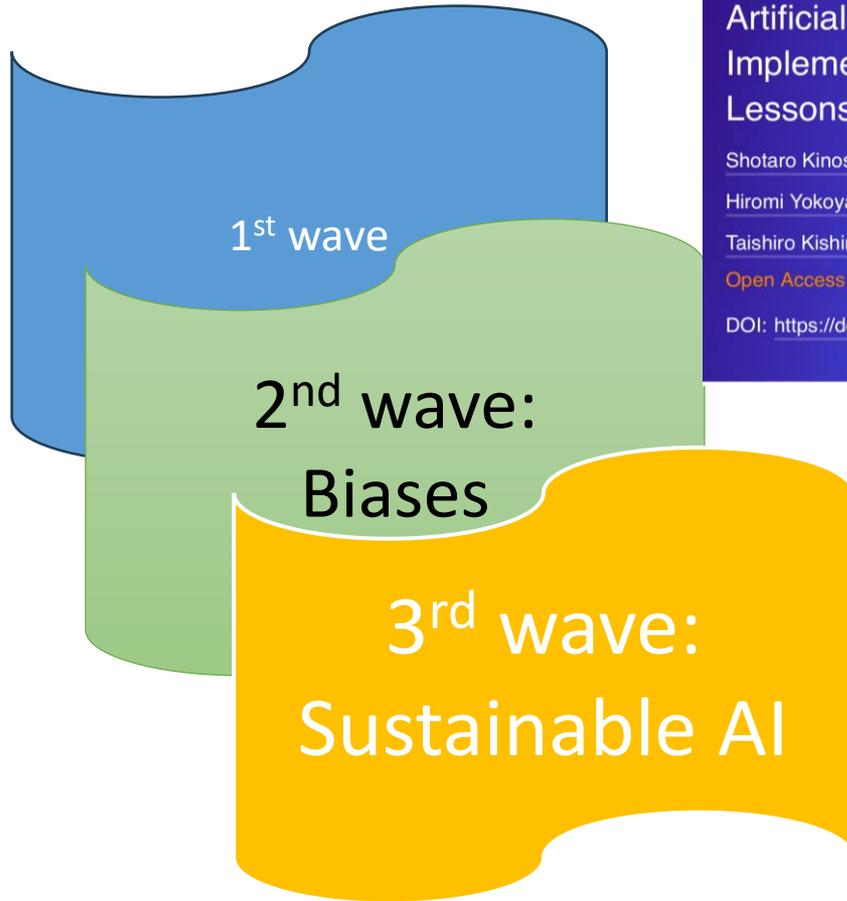
企業で科学論や哲学者を戦略的に組み入れる可能性

・ 日経新聞に上記内容を寄稿

2023. 5. 9 [AI倫理の指針「AI8原則」の順守を 人工知能と社会](#)

・ Nature [Correspondence "Large language model is a flagship for Japan" Shoraro Kinoshita and Hiromi Yokoyama](#)

AI倫理は第3の波へ -「サステイナブルAI」



LETTER TO THE EDITOR | VOLUME 2, ISSUE 2, P258,
JUNE 2024

Cautions and Considerations in Artificial Intelligence Implementation for Child Abuse: Lessons from Japan

Shotaro Kinoshita, MD, PhD •
Hiromi Yokoyama, PhD •
Taishiro Kishimoto, MD, PhD

[Open Access](#) • Published: April 15, 2024 •
DOI: <https://doi.org/10.1016/j.mcpdig.2024.04.005>

nature reviews physics Search Log in

Content ▾ About ▾ Publish ▾

[viewpoint](#) > article

Viewpoint | Published: 04 March 2024

Generative AI and science communication in the physical sciences

[Sibusiso Biyela](#) , [Kanta Dihal](#) , ... [Hiromi M. Yokoyama](#) [+ Show authors](#)

[Nature Reviews Physics](#) **6**, 162–165 (2024) | [Cite this article](#)

4345 Accesses | **1** Citations | **16** Altmetric | [Metrics](#)

Attitudes Toward Facial Analysis AI: A Cross-National Study Comparing Argentina, Kenya, Japan, and the USA

Chiara Ullstein
chiara.ullstein@tum.de
Technical University of Munich,
Chair of Cyber Trust
Munich, Germany

Yuko Ikkatai
y.ikkatai@staff.kanazawa-u.ac.jp
Kanazawa University
Kanazawa, Japan

Brian Mboya
brian.mboya@protonmail.com
Dedan Kimathi University of
Technology, Computer Science
Nairobi, Kenya

Severin Engelmann
severin.engelmann@cornell.edu
Cornell Tech,
Digital Life Initiative
New York City, USA

Naira-Paola Arnez-Jordan
naira.arnez@tum.de
Technical University of Munich
Munich, Germany

Shuichiro Higuma
shuichiro-higuma@g.ecc.u-
tokyo.ac.jp
The University of Tokyo
Tokyo, Japan

Orestis Papakyriakopoulos
orestis.p@tum.de
Technical University of Munich,
Professorship of Societal Computing
Munich, Germany

Rose Caleno
calenorose@gmail.com
Nairobi, Kenya

Tilman Hartwig
Tilman.Hartwig@uba.de
Umweltbundesamt,
AI Lab
Leipzig, Germany

Hiromi Yokoyama
hiromi.yokoyama@ipmu.jp
The University of Tokyo,
Kavli IPMU, CD3
Tokyo, Japan

Jens Grossklags
jens.grossklags@in.tum.de
Technical University of Munich,
Chair of Cyber Trust
Munich, Germany