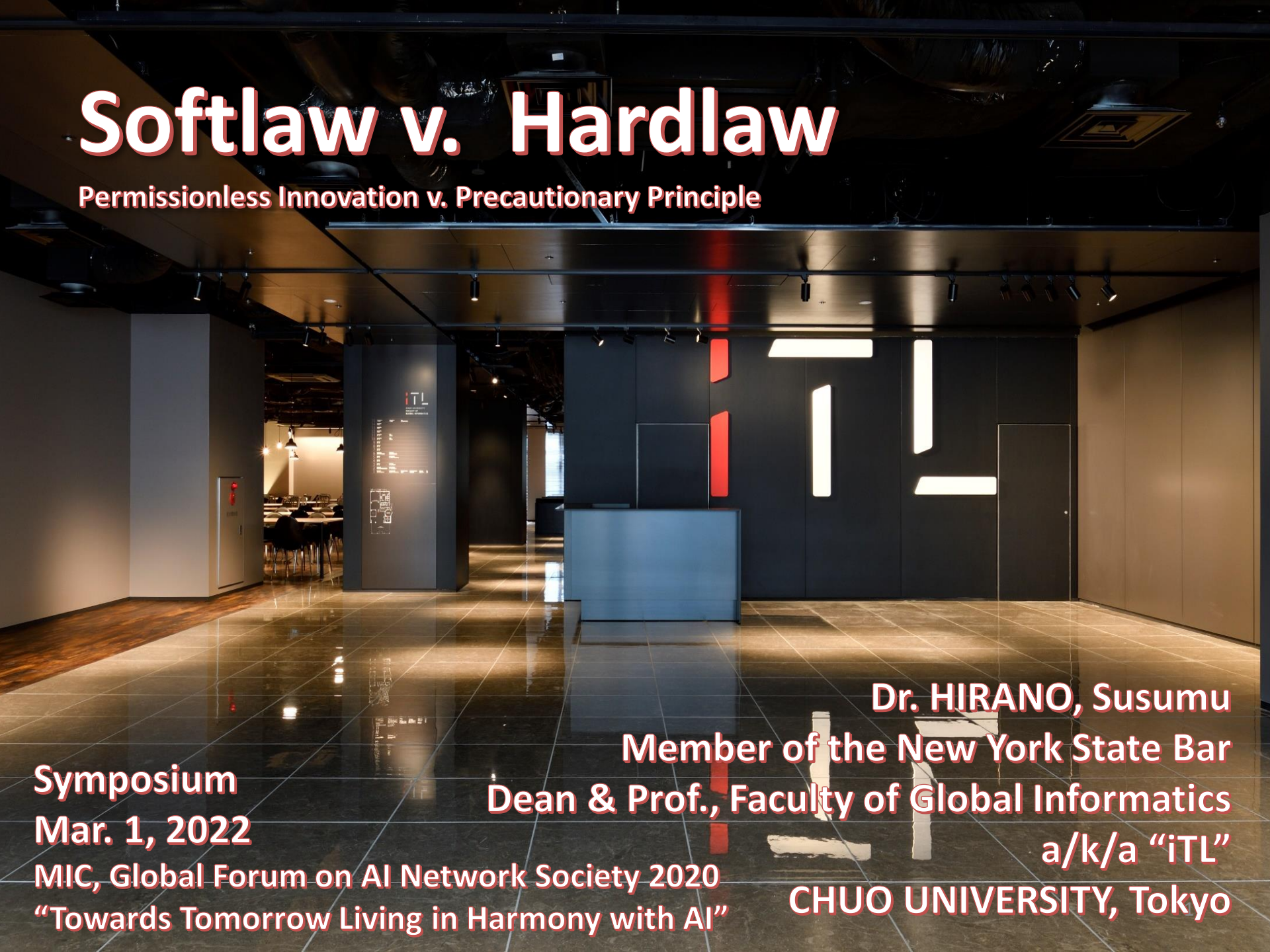


Softlaw v. Hardlaw

Permissionless Innovation v. Precautionary Principle

Dr. HIRANO, Susumu
Member of the New York State Bar
Dean & Prof., Faculty of Global Informatics
a/k/a “iTL”
CHUO UNIVERSITY, Tokyo

Symposium
Mar. 1, 2022
MIC, Global Forum on AI Network Society 2020
“Towards Tomorrow Living in Harmony with AI”

Self-Introduction

Professor and Dean, Faculty of Global Informatics, Chuo University

Doctor (Pol'y Stud.) 2007, Chuo University (Tokyo, Japan); LL.M. 1990, Cornell University (Ithaca, New York); and LL.B. 1984, Chuo University (Tokyo, Japan). Admitted to New York State Bar in and after Apr. 1991. Susumu Hirano has been professor of law (tenured) since 2004 at Chuo University. He was the founder of the Faculty of Global Informatics and has been Dean thereof since its foundation in Apr. 2019. Before that he had been Dean, Graduate School of Policy Studies since 2013 till 2019. Before he became the tenured professor at Chuo University, he had been General Counsel, Legal Department, NTT DoCoMo, Inc. (Tokyo, Japan). He has been active in the field of ELSI of AI including, but not limited to: a member of AIGO (AI expert Group at OECD)(Paris, France); a member of Council for Social Principles of Human-centric AI, Cabinet Office, Government of Japan; and Chairperson, Committee on AI R&D Principles, Conference toward AI Network Society, Ministry of Internal Affairs and Communications, Japan (MIC).

He wrote, inter alia, ROBOT LAW (Kobundo, 2nd ed., 2019, in Japanese); AMERICAN CONTRACTS (Chuo Univ. Press, 2009, in Japanese); AMERICAN TORTS (Chuo Univ. Press, 2006, in Japanese); and ELECTRONIC COMMERCE AND CYBERLAW (NTT Press, 1999, in Japanese).

"AIGO:" Artificial Intelligence expert Group at OECD in Paris in Sept. 2018.

Japan's Contribution toward Global Standards

US

Permissionless Innovation

versus

EU

Precautionary Principle

JP

Non-binding Approach = Soft Law

Non-binding norms

See SUSUMU HIRANO, ROBOT LAW, 267-68 (*Kobundo* 2nd ed. 2019 in Japanese).

AI Principles ／ AI Guidelines:

Japan Took the Initiative to Build Global Standards

G20 **June 2019**

- **G20 AI Principles** / Ministerial Meeting on Trade and Digital Economy

OECD BETTER POLICIES FOR BETTER LIVES

OECD **Sept. 2018 to May 2019**

- **OECD AI Principles** / Council Recommendation on AI (May 2019)
- **AIGO** Sep. 2018 to Feb. 2019 ✓

内閣府 Cabinet Office

Cabinet Office **Apr. 2018 to Mar. 2019**

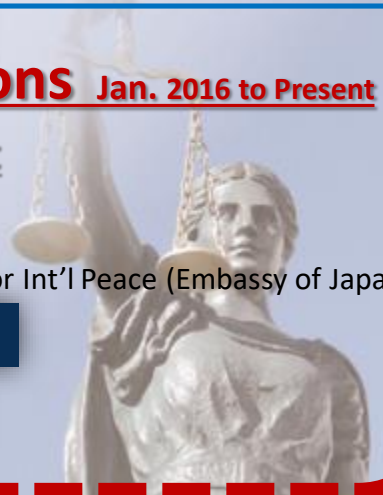
- **Council of Principles of Human-centric AI**

総務省 MIC Ministry of Internal Affairs and Communications

✓ **MIC: Ministry of Internal Affairs & Communications** **Jan. 2016 to Present**

- The Conference toward AI Network Society **Oct. 2016 to Present**
  - Conference on AI: Intelligent Machines, Smart Policies, Co-sponsored by MIC and OECD
  - Forum toward AI Network Society (Int'l Symposium) in Tokyo, Sponsored by MIC.
  - Conference on Artificial Intelligence and US-Japan Alliance Engagement, Carnegie Endowment for Int'l Peace (Embassy of Japan in the U.S.A. and MIC)
  - Technology Foresight Forum 2016 on AI, OECD
  - G7 ICT Ministers' Meeting in Takamatsu, Kagawa
- Conference on Networking among AIs **Jan. to June 2016**

CARNEGIE ENDOWMENT FOR INTERNATIONAL PEACE



OECD・AI専門家会合

AI expert *G*roup at the *O*ECD （AIGO：エイゴ／エイ・アイ・ゴー）

AIGO’s 1st meeting at OECD in Paris, Sept. 24-25, 2018

AIGO’s 2nd meeting at OECD in Paris, Nov. 12, 2018

OECD.AI Policy Observatory

OECD.org | Going Digital Toolkit | E

Experts & blog | AI Principles | Policy areas | Trends & data | Countries & tools | About

Home > OECD AI Principles overview

OECD AI Principles overview

The OECD AI Principles promote use of AI that is innovative and trustworthy and that respects human rights and democratic values. Adopted in May 2019, they set standards for AI that are practical and flexible enough to stand the test of time.

Values-based principles

- Inclusive growth, sustainable development and well-being
- Human-centred values and fairness
- Transparency and explainability
- Robustness, security and safety
- Accountability

OECD AI, https://oecd.ai/en/ai-principles (last visited Feb. 18, 2022).

G20_AI Principles

G20 Ministerial Statement on Trade and Digital Economy

1. We, the G20 Trade Ministers and Digital Economy Ministers, met on 8 and 9 June 2019 in Tsukuba City, Ibaraki Prefecture, Japan, under the chairmanship of H.E. Mr. Hiroshige Seko, Minister of Economy, Trade and Industry, H.E. Mr. Masatoshi Ishida, Minister for Internal Affairs and Communications, and H.E. Mr. Taro Kono, Minister for Foreign Affairs, of the Government of Japan, to further strengthen G20 trade and digital economic policy cooperation.

2. The G20 Ministerial Meeting on Trade and Digital Economy gathered all G20 members as well

. . .

19. At the same time, we also recognize that AI, like other emerging technologies, may present societal challenges, including the transitions in the labor market, privacy, security, ethical issues, new digital divides and the need for AI capacity building. To foster public trust and confidence in AI technologies and fully realize their potential, we are committed to a human-

3

centered approach to AI, guided by the G20 AI Principles drawn from the OECD Recommendation on AI, which are attached in Annex and are non-binding. This Annex includes the following principles of "inclusive growth, sustainable development and well-being", "human-centered values and fairness", "transparency and explainability", "robustness, security and safety" and "accountability". The Annex also offers guidance for consideration by policy makers with the purpose of maximizing and sharing the benefits from AI, while minimizing the risks and concerns, with special attention to international cooperation and inclusion of developing countries and underrepresented populations.

20. In pursuing human-centered AI, G20 members recognize the need to continue to promote the

ANNEX

G20 AI Principles [2]

The G20 supports the Principles for responsible stewardship of Trustworthy AI in Section 1 and takes note of the Recommendations in Section 2.

Section 1: Principles for responsible stewardship of trustworthy AI

1.1. Inclusive growth, sustainable development and well-being

Stakeholders should proactively engage in responsible stewardship of trustworthy AI in pursuit of beneficial outcomes for people and the planet, such as augmenting human capabilities and enhancing creativity, advancing inclusion of underrepresented populations, reducing economic, social, gender and other inequalities, and protecting natural environments, thus invigorating inclusive growth, sustainable development and well-being.

1.2. Human-centered values and fairness

a) AI actors should respect the rule of law, human rights and democratic values, throughout the AI system lifecycle. These include freedom, dignity and autonomy, privacy and data protection, non-discrimination and equality, diversity, fairness, social justice, and internationally recognized labor rights.

b) To this end, AI actors should implement mechanisms and safeguards, such as capacity for human determination, that are appropriate to the context and consistent with the state of art.

1.3. Transparency and explainability

AI Actors should commit to transparency and responsible disclosure regarding AI systems. To this end, they should provide meaningful information, appropriate to the context, and consistent with the state of art:

i. to foster a general understanding of AI systems;

ii. to make stakeholders aware of their interactions with AI systems, including in the workplace;

iii. to enable those affected by an AI system to understand the outcome; and,

iv. to enable those adversely affected by an AI system to challenge its outcome based on plain and easy-to-understand information on the factors, and the logic that served as the basis for the prediction, recommendation or decision.

1.4. Robustness, security and safety

a) AI systems should be robust, secure and safe throughout their entire lifecycle so that, in

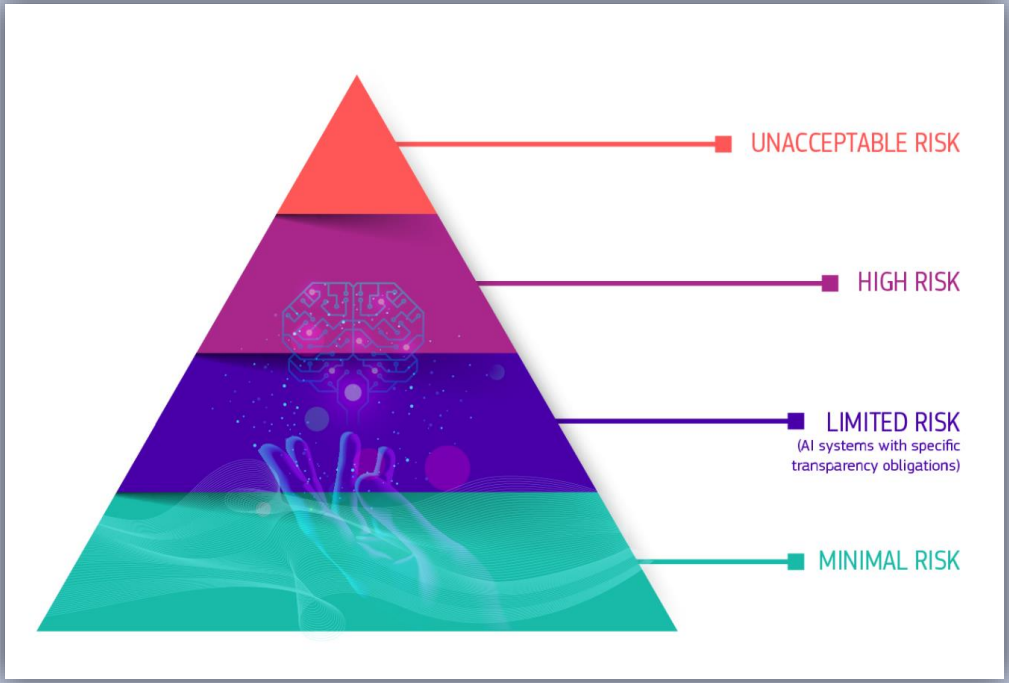
[2] This Annex draws from the OECD principles and recommendations.

11

"drawn from～:" ～から徴用した、～を利用した

G20貿易・デジタル経済大臣会合閣僚声明（英文）, http://www.soumu.go.jp/main_content/000625715.pdf (last visited Nov. 8, 2019)(emphasis added).

EU's Hardlaw Proposal: Artificial Intelligence Act



☑ **Risk-Based Control**

European Commission, Excellence and trust in artificial intelligence https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/excellence-trust-artificial-intelligence#building-trust-through-the-first-ever-legal-framework-on-ai (last visited May 5, 2021).

EU's Hardlaw Proposal: Artificial Intelligence Act

Tiers	Categories	*E.g.*:	Authorities
Tier 1	Prohibited AI Practices / Un-acceptable risk	*E.g.*: ✓ **subliminal techniques**; ✓ an AI system that **exploits vulnerabilities** to distort materially distort their behaviors ; ✓ AI-based **social scoring** for general purposes done by public authorities; and ✓ **'real time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement** unless certain limited exceptions apply.	*See* Title II: Art. 5
Tier 2	High-Risk AI Systems	*E.g.*: ✓ Safety component of products that are subject to third party ex-ante conformity assessment; or ✓ Stand-alone AI systems, with mainly fundamental rights implications, listed in ANNEX III.	*See* Title III
Tier 3	Transparency Obligations for Certain AI Systems / Specific risks of manipulation	Transparency obligations will apply for systems that: (i) interact with humans; (ii) are used to detect emotions or determine association with (social) categories based on biometric data; or (iii) generate or manipulate content ('deep fakes').	*See* Title IV
Tier 4	Codes of Conduct	"The Commission and the Member States shall encourage and facilitate the drawing up of codes of conduct intended to foster the **voluntary application** to **AI systems other than high-risk AI systems** of the requirements set out in Title III, Chapter 2. . . ." / "Codes of conduct may be drawn up by individual providers of AI systems or by organisations representing them or by both, including with the involvement **of users and any interested stakeholders** and their representative organisations."	*See* Title IX, Art. 69

European Parliament, Committee on Legal Affairs, a Civil Liability Regime for Artificial Intelligence

Committee on Legal Affairs, Report with Recommendations to the Commission on a Civil Liability Regime for Artificial Intelligence (2020/2014 (INL)), Oct. 5, 2020.

- 「人工知能の為の民事賠償責任レジーム」
- 厳格責任、過失の推認、等の立法提案を、欧州議会が決議（2020年10月）

high-risk AI operatorの厳格責任

Article 4

Strict liability for high-risk AI-systems

1. The operator of a high-risk AI-system shall be strictly liable for any harm or damage that was caused by a physical or virtual activity, device or process driven by that AI-system.

2. All high-risk AI-systems and all critical sectors where they are used shall be listed in the Annex to this Regulation. The Commission is empowered to adopt delegated acts in accordance with Article 13, to amend that exhaustive list, by:

(a) including new types of high-risk AI-systems and critical sectors in which they are deployed;

(b) deleting types of AI-systems that can no longer be considered to pose a high risk; and/or

(c) changing the critical sectors for existing high-risk AI-systems.

Any delegated act amending the Annex shall come into force six months after its adoption. When determining new high-risk AI-systems and/or critical sectors to be inserted by means of delegated acts in the Annex, the Commission shall take full account of the criteria set out in this Regulation, in particular those referred to in Article 3(c).

3. Operators of high-risk AI-systems shall not be able to exonerate themselves from liability by arguing that they acted with due diligence or that the harm or damage was caused by an autonomous activity, device or process driven by their AI-system. Operators shall not be held liable if the harm or damage was caused by force majeure.

4. The frontend operator of a high-risk AI-system shall ensure that operations of that AI-system are covered by liability insurance that is adequate in relation to the amounts and extent of compensation provided for in Articles 5 and 6 of this Regulation. The backend operator shall ensure that its services are covered by business liability or product liability insurance that is adequate in relation to the amounts and extent of compensation provided for in Article 5 and 6 of this Regulation. If compulsory insurance regimes of the frontend or backend operator already in force pursuant to other Union or national law or existing voluntary corporate insurance funds are considered to cover the operation of the AI-system or the provided service, the obligation to take out insurance for the AI-system or the provided service pursuant to this Regulation shall be deemed fulfilled, as long as the relevant existing compulsory insurance or the voluntary corporate insurance funds cover the amounts and the extent of compensation provided for in Articles 5 and 6 of this Regulation.

5. This Regulation shall prevail over national liability regimes in the event of conflicting strict liability classification of AI-systems.

Civil Liability Regime for Artificial Intelligence, *supra*.

Softlaw v. Hardlaw

- Emerging technologies are moving targets.
- They might become obsolete soon so that the hardlaw therefor could become inoperative, too.
- Hardlaw in an early stage of a certain emerging technology might become over-deterrent and hinder innovation which could save many people.
- Thus, for current AI systems, flexible softlaw might be more appropriate than rigid hardlaw.
- However, softlaw's drawback is lack of enforceablity.
- **Anyway, EU's hardlaw proposals are at least worth studying.**
- When we study them, we should consider, including, but not limited to, reasonableness of the rules proposed by EU, and acceptability of the values esteemed by EU.

Reasonableness: Benefit/Probability Neglect

- "Fear" hinders reasonable thinking.
- Unscientific precautions are often taken due to fear.
  - *E.g.*, sea or agricultural food products caught or produced in *Fukushima* are refused to be imported in some countries, even though they are not radio active.
- "Fear" makes the people *ignore* the probability element in the expected accident costs.
- "Fear," also, makes the people *ignore* the benefits of emerging technologies.
- Is the hardlaw proposed by EU reasonable in each and every aspect?

See generally Cass R. Sunstein, *Beyond the Precautionary Principle*, 151 U. PA. L. REV. 1003 (2003).

"Fear"

So, first of all, let me assert my firm belief that the only thing we have to fear is fear itself—nameless, unreasoning, unjustified terror which paralyzes needed efforts to convert retreat into advance.

Franklin Roosevelt's First Inaugural Address

Risk Perceptions Differ

- Different regions have different perceptions of risks.
  - *E.g.*, Nuclear risk might be perceived as one of the highest risks in Japan, while in EU and US risk ranking thereof might be lower.

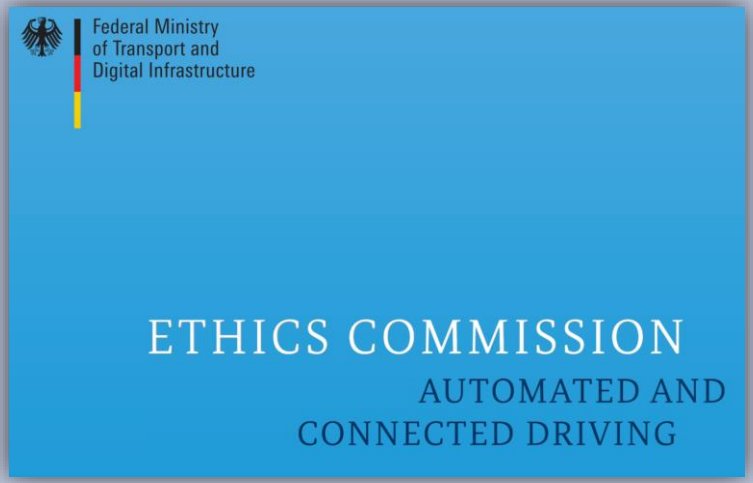
- Japan does not have to share the European values in each and every aspect.

E.g., Crush Optimization or Derivative Trolley Problems

Germany seems to adopt the so-called absolute deontology.

9.

In the event of unavoidable accident situations, any distinction based on personal features (age, gender, physical or mental constitution) is strictly prohibited. It is also prohibited to offset victims against one another. General programming to reduce the number of personal injuries may be justifiable. Those parties involved in the generation of mobility risks must not sacrifice non-involved parties. (emphasis added)



Federal Ministry of Transportation and Digital Infrastructure, Ethics Commission, Report, *Automated and Connected Driving*, June 2017, at 11, https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile (last visited Nov. 22, 2021)(emphasis added).

E.g., Crush Optimization or Derivative Trolley Problems

1.6 No selection of humans, no offsetting of victims, but principle of damage minimization

The modern constitutional state only opts for absolute prohibitions in borderline cases, such as the ban on torture relating to persons in state custody.[4] Regardless of the consequences, an act is mandated or prohibited absolutely because it is intrinsically already incompatible with the constitutive values of the constitutional order. Here, there is, exceptionally, no trade-off, which is per se a feature of any morally based legal regime. The Federal Constitutional Court's judgment on the Aviation Security Act [5] also follows this ethical line of appraisal, with the verdict that the sacrifice of innocent people in favour of other potential victims is impermissible, because the innocent parties would be degraded to mere instrument and deprived of the quality as a subject. This position is not without controversy, either in constitutional law [6] or ethically [7], but it should be observed by lawmakers. (emphasis added)

Id. at 18 (emphasis added).

E.g., Crush Optimization or Derivative Trolley Problems

However, the Ethics Commission refuses to infer from this that the lives of humans can be „offset" against those of other humans in emergency situations so that it could be permissible to sacrifice one person in order to save several others. It classifies the killing of or the infliction of serious injuries on persons by autonomous vehicles systems as being wrong without exception. Thus, even in an emergency, human lives must not be „offset" against each other. According to this position, the individual is to be regarded as „sacrosanct". No obligations of solidarity must be imposed on individuals requiring them to sacrifice themselves for others, even if this is the only way to save other people.

A different decision may have to be taken if several lives are already imminently threatened and the only thing that matters is saving as many innocent people as possible. In situations of this kind, it would appear reasonable to demand that the course of action to be chosen is that which costs as few human lives as possible. Here, the Commission has not yet been able to bring its discussions to a satisfactory end, nor has it been able to reach a consensus in every respect. It thus suggests that in-depth studies be conducted. [8] (emphasis added)

Id. at 18 (emphasis added).

E.g., Crush Optimization or Derivative Trolley Problems

『NATURE』誌のMoral Machine Experiment論文も、2017 German Ethics Commission Reportと大衆嗜好との齟齬を以下のように指摘:


ARTICLE

https://doi.org/10.1038/s41586-018-0637-6

The Moral Machine experiment

Edmond Awad[1], Sohan Dsouza[1], Richard Kim[1], Jonathan Schulz[2], Joseph Henrich[2], Azim Shariff[3]*, Jean-François Bonnefon[4]* & Iyad Rahwan[1,5]*

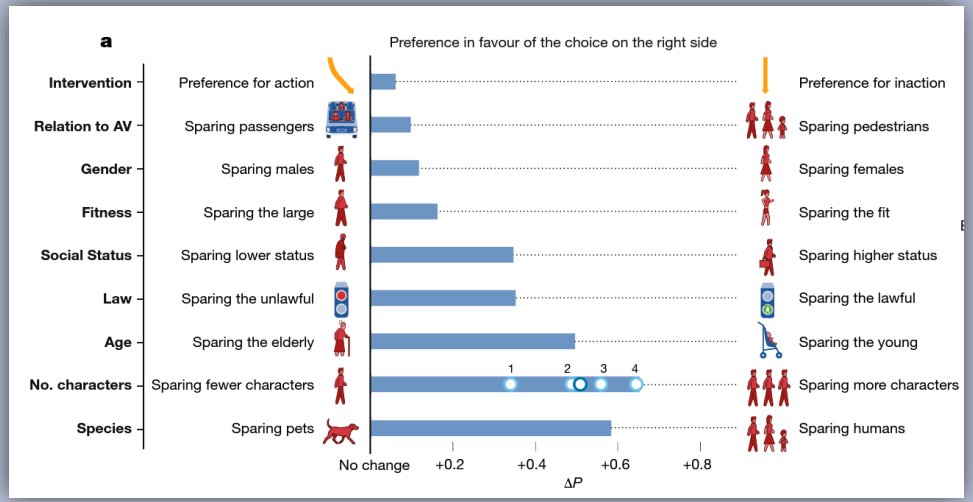
Consider, as a case in point, the ethical rules proposed in 2017 by the German Ethics Commission on Automated and Connected Driving []. This report represents the first and only attempt so far to provide official guidelines for the ethical choices of autonomous vehicles. As such, it provides an important context for interpreting our findings and their relevance to other countries that might attempt to follow the German example in the future. . . . . On the other hand, German Ethical Rule number 9 does not take a clear stance on whether and when autonomous vehicles should be programmed to sacrifice the few to spare the many, but leaves this possibility open: it is important, thus, to know that there would be strong public agreement with such programming, even if it is not mandated through regulation.

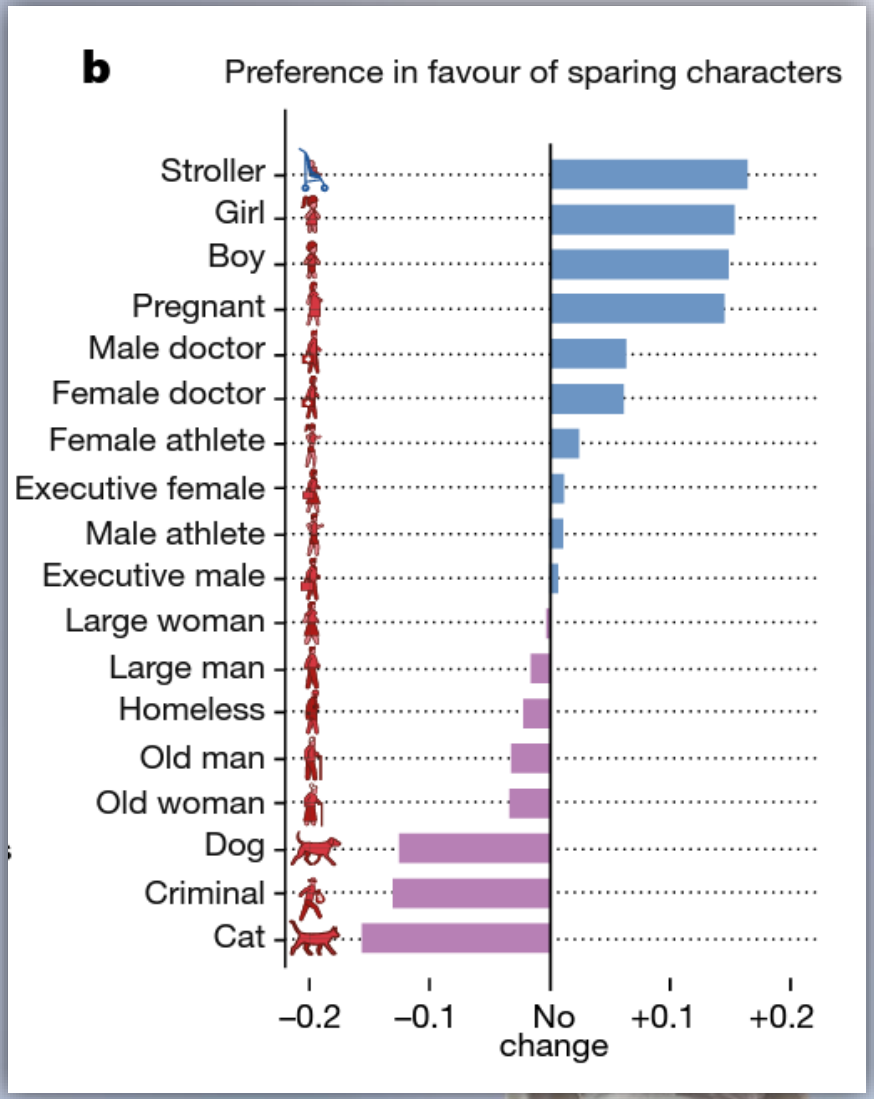
Jean François Bonnefon et al., *The Moral Machine Experiment*, 563 NATURE 59, 60 (Nov. 1, 2018)(emphasis added).

E.g., Crush Optimization or Derivative Trolley Problems

By contrast, German Ethical Rule number 9 also states that any distinction based on personal features, such as age, should be prohibited. This clearly clashes with the strong preference for sparing the young (such as children) that is assessed through the Moral Machine (see Fig. 2b for a stark illustration: the four most spared characters are the baby, the little girl, the little boy, and the pregnant woman). This does not mean that policymakers should necessarily go with public opinion and allow autonomous vehicles to preferentially spare children, or, for that matter, women over men, athletes over overweight persons, or executives over homeless persons—for all of which we see weaker but clear effects. But given the strong preference for sparing children, policymakers must be aware of a dual challenge if they decide not to give a special status to children: the challenge of explaining the rationale for such a decision, and the challenge of handling the strong backlash that will inevitably occur the day an autonomous vehicle sacrifices children in a dilemma situation. *Id.* at 60 (emphasis added).

E.g., Crush Optimization or Derivative Trolley Problems





Id. at 61 Figs. 2a & 2b.

E.g., Crush Optimization or Derivative Trolley Problems

Fig. 2 | Global preferences. a, AMCE for each preference. In each row, ΔP is the difference between the probability of sparing characters possessing the attribute on the right, and the probability of sparing characters possessing the attribute on the left, aggregated over all other attributes. For example, for the attribute age, the probability of sparing young characters is 0.49 (s.e. = 0.0008) greater than the probability of sparing older characters. The 95% confidence intervals of the means are omitted owing to their insignificant width, given the sample size (n = 35.2 million). For the number of characters (No. characters), effect sizes are shown for each number of additional characters (1 to 4; n_1 = 1.52 million, n_2 = 1.52 million, n_3 = 1.52 million, n_4 = 1.53 million); the effect size for two additional characters overlaps with the mean effect of the attribute. AV, autonomous vehicle. **b**, Relative advantage or penalty for each character, compared to an adult man or woman. For each character, ΔP is the difference the between the probability of sparing this character (when presented alone) and the probability of sparing one adult man or woman (n = 1 million). For example, the probability of sparing a girl is 0.15 (s.e. = 0.003) higher than the probability of sparing an adult man or woman.

Id. at 61.

CBA: Cost-Benefit Analysis

- **Hand Formula**

  B<PL → Negligence

  B: Burden / P: Probability / L: Loss

- **Untaken Precautions / RAD(*1)**

  (*1) a Reasonable Alternative Design

  - **Comparison of a RAD with the subject design**

Thank You

国際情報学部 | 中央大学 (chuo-u.ac.jp)