

## モデレータ情報共有

- ✓ 「ナッジ」について
  - ✓ アーキテクチャ (choice architecture) による影響
  - ✓ 「Code 理論」
- 
- ナッジの定義

A nudge is "any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives."

Michal Lavi, *Evil Nudge*, 21 VAND. J. ENT. & TECH. L. 1, 1  
(2018) (emphasis added).

- "libertarian paternalism" / "asymmetric paternalism"
- *status quo bias* → default + opt-out
- 悪用も在り得る → "evil nudge" = 福祉を増進しないナッジ *Id.* at 10.
- 諸刃の剣。→ AI と同じかも?!

### EU AI 法案

- サブリミナルなどの人を操作する AI システムへの規制を明言

#### *Article 5*

1. The following artificial intelligence practices shall be prohibited:

- (a) the placing on the market, putting into service or use of an AI system that deploys **subliminal techniques beyond a person's consciousness** in order to materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm;

- (b) the placing on the market, putting into service or use of an AI system that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, in order to materially distort the behaviour of a person pertaining to that group in a manner that causes or is likely to cause that person or another person physical or psychological harm;

#### 日本の内閣府「人間中心の AI 社会原則」会議

- 過度の依存や悪用による意思の操作を、望ましくない事象であるとしている。

##### (1) 人間中心の原則

AI の利用は、憲法及び国際的な規範の保障する基本的人権を侵すものであってはならない。

AI は、人々の能力を拡張し、多様な人々の多様な幸せの追求を可能とするために開発され、社会に展開され、活用されるべきである。AI が活用される社会において、人々が AI に過度に依存したり、AI を悪用して人の意思決定を操作したりすることのないよう、我々は、リテラシー教育や適正な利用の促進などのための適切な仕組みを導入することが望ましい。

- 「自由意思」とは何か？
- ヒトのあらゆる意思・判断は、所詮は何らかの影響を受けているのであるから、そもそも「自由意思」など存在しないのでないか？
  - 最近の有名な例：E.g. 〈フィルター・バブル〉
  - 法律の例) E.g. 〈クーリング・オフ〉 → 本人の利益に成るか否かが分かれ目かも。
  - 消費者保護関連諸法の世界では、以前から、誤認を許さない法が存在。E.g., 景品表示法は**優良誤認**を違法と規定。
- 要は程度の問題かも。
- 分かっていたつもりでも、実は私達は自分のことを（人間自身のことを）何も理解していなかった— AI やロボットに係る法分野（ロボット法）は、人間自身のことを知ること・人間自身の探求である、といわれている。