

リスク社会の新たな局面

大阪大学ELSIセンターシンポジウム
「科学技術イノベーションと倫理・法・社会」

久木田水生

名古屋大学大学院情報学研究科
minao.kukita@i.nagoya-u.ac.jp

概要

- 人工知能が社会に蔓延することはリスク要因である。
- 一方で人工知能は優れたリスク分析、リスク管理の道具でもある。
- 特にビッグデータに基づく機械学習は人間を対象としたリスク分析のための画期的なツールである。
- しかしリスク分析ツールとしての人工知能が普及することにもまた固有のリスクがある。
- 本発表では特に人工知能の「メディア」としての特性に注目して、そのリスクについて考察する。

リスク社会

- 社会学者、ウルリッヒ・ベックの言葉。
- 経済、産業、テクノロジーの進歩、グローバル化の進展によって、何らかの大規模な惨事の可能性が前提として織り込まれた社会。
- 典型的にはテロ、金融危機、自然破壊など。

AIリスク社会

- 自律的に稼働する膨大な数の多種多様なAI機器、ソフトウェアが社会の隅々で使われ、複雑なネットワークを形成して、予測不可能な大惨事を引き起こす可能性がぬぐえなくなった社会。
- ウォラック&アレン『ロボットに倫理を教える』の冒頭では、ちょっとしたきっかけで人工知能が連鎖的に誤作動を起こし、カタストロフ的な帰結に至るシナリオが描写されている。

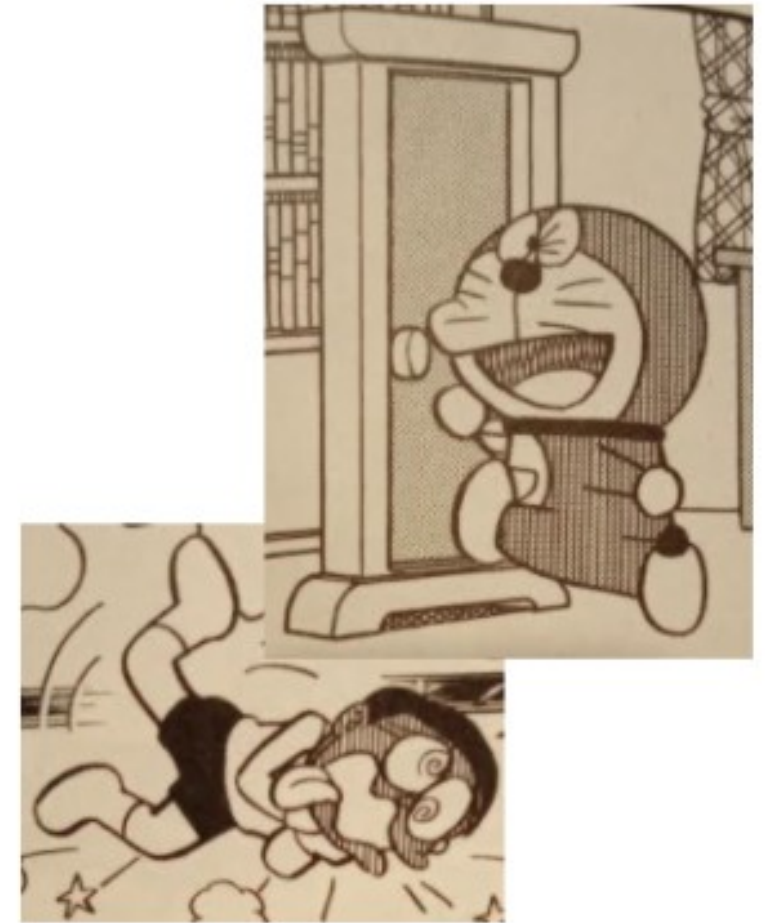


AIリスク社会

- しかしAIのリスクはこのようなカタストロフ的な、一時的な大災害だけではない。
- むしろ目立たないが蔓延し、常態化し、人々がそれに対し無感覚になる危害の方が危険かもしれない。

殺人どこでもドア

- 通常のどこでもドアと同じように、行きたい場所どこにでも一瞬で連れて行ってくれる。しかしそのドアをくぐった人は一定の確率で様々な程度の怪我をして、運が悪ければ死に至ることもある。さらにはどういう訳か誰かがドアを利用すると一定の確率で無関係の人間が危害を受ける。
- 日本中の誰もが日常の移動にこのドアを使うようになると、年間数千人程度の命が犠牲になり、数十万人の人間が様々な程度の怪我をすると推定される。
- このようなドアが発明されたら社会はそれを受け入れるだろうか？



昨年度の交通事故発生状況

1 交通事故発生状況

表1-1 交通事故発生状況

	12月末	前年同期比	
		増減数	増減率
交通事故件数	309,178	-72,059	-18.9
うち死亡事故件数	2,784	-349	-11.1
死者数	2,839	-376	-11.7
負傷者数	369,476	-92,299	-20.0

「令和2年中の交通死亡事故の発生状況及び道路交通法違反取締り状況等について」
<https://www.e-stat.go.jp/>

ビッグデータと機械学習

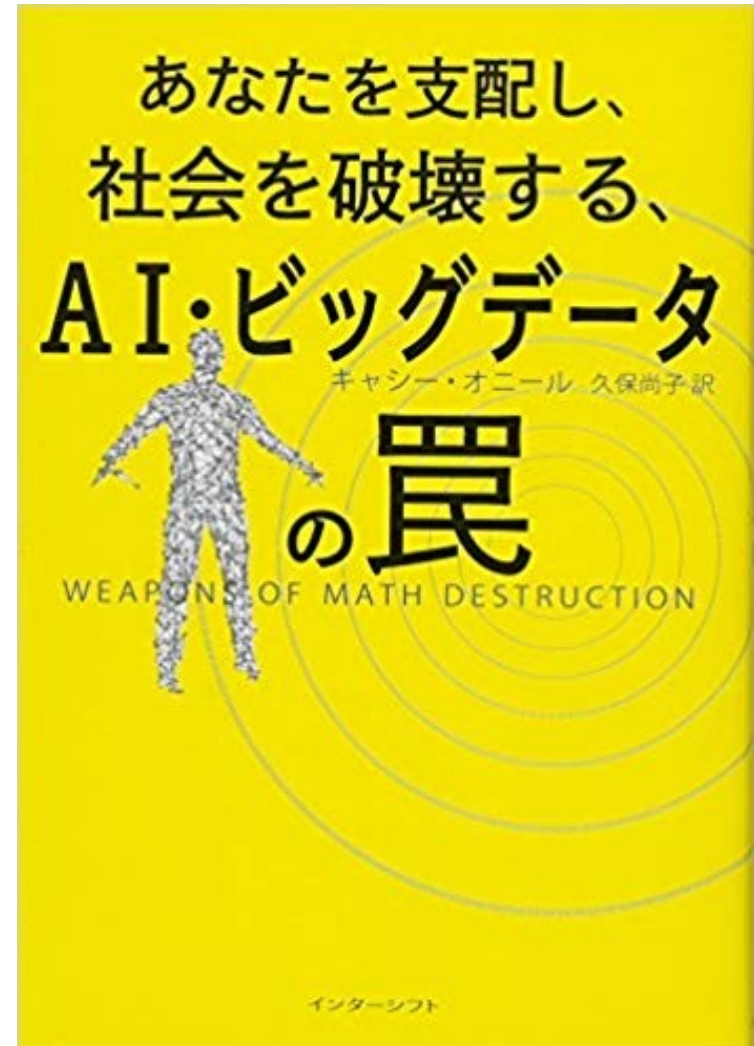
- 人工知能と呼ばれるテクノロジーは多様だが、現在、最も活用され、莫大な利益を生み出しているのは、ビッグデータに基づいて学習を行い、人々の属性、能力、性格、健康状態、行動傾向などを予測する人工知能。
- このような人工知能は、広告、商品推薦、ニュースフィード、オンラインショッピングの価格決定、保険料金の決定、犯罪発生予測、医療診断、信用スコアなどに使われている。

AIリスク管理社会

- 人工知能、特にビッグデータに基づく機械学習は優れた「リスク分析」の道具である。
- 従来、人間の属性や性格や能力や行動を正確に推測・予測するのは困難であった。
- スマートフォンなどのデバイスを通じて得られるビッグデータと、それに基づいて学習を行なう人工知能は、こういったことについて、以前よりはるかに正確な予測を可能にした。
- そのことによってデータと人工知能を活用するIT企業や政府は、リスクの少ない意思決定を行なうことができるようになる。

AIリスク管理のリスク

- こういった人工知能は巨大IT企業に莫大な利益をもたらしている一方で、しばしば社会的弱者をさらに不利な状況に陥れ、差別や格差を助長することが指摘されている。



「AIだから公平、客観的」はペテン師の言葉

- ビッグデータに基づく機械学習は必ずしも公平、客観的なものではなく、それを設計した人間の先入観や、データの持つ偏りを反映する。
- さらにディープラーニングなどを使った機械学習ではその判断の理由が人間には理解できないことがしばしばであり、ブラックボックスになる。
- それをいいことに、差別的な判断を「人工知能がそう判断した」というような言い訳する人間、いい加減に作ったシステムを売り込む人間が現れる。
- 「人工知能だから公平」、「人工知能だから客観的」という人間はペテン師か、人工知能をよく分かっていないかのどちらか。

なぜペテンを信じるのか？

- 人は信じたいものを信じる。
- ビッグデータと人工知能が、複雑な世界の複雑な問題に、簡単な「解決」をもたらしてくれることを人々は期待しているのかもしれない。
- あるいは人工知能が自分たちの持っている偏見や先入観を確証してくれるからこそ、それを信じるのかもしれない。
- ただ単に「数学」、「アルゴリズム」、「人工知能」といった言葉に幻惑されるのかもしれない。

「メディアはメッセージ」

- マーシャル・マクルーハンの言葉。
- 伝達される情報の形式や様態（モダリティ）、あるいは私たちがその情報を消費する仕方は各々のメディアの持つ技術的特性によって規定される。
- 同じ内容のメッセージでも、どのメディアによって伝えられるかによって、受け取った人間に与える効果は異なる。
- どのメディアを使うかということは、メッセージの内容以上に重要な意味を持つ。

「メディアはメッセージ」

- 例えば新聞ならば伝えられる情報は文字だけ、ラジオならば音声だけである。
- 新聞は私たちが情報を消費する間、そこに注意を集中することを要求するのに対して、ラジオは私たちが他の仕事をしながら（運転しながら、本を読みながら、食事をしながら、ジョギングをしながら）聞き流すような消費の仕方を可能にする。
- また新聞の情報は保存がきくのに対してラジオはそうではない、音声は文字に比べて感情を喚起しやすいなどなどの違いがある。

メディアとしての人工知能

- 伝統的なメディアが基本的に事実を伝えるのに対して、人工知能は一般にデータに基づいた確率的な推測を伝える。
- もちろん新聞やニュースも推測を伝えるが、それは人間が行った推測を伝えているだけである。
- それに対して人工知能は機械が行なった推測を伝える。
- しかもそれは個々の私人（あるいは私人の集団）の行動や性格や能力や思考についての確率的な推測である。

人工知能のメッセージ

- 人間は様々なデバイス、アプリケーションを通じて取得されるデータと、そこから確率的に推測される属性の集まりとして理解できる。
- 他者はあなたにとってリスクであり、そのリスクは前もって見積り、回避することができる。

他者をリスクとして扱うことの弊害

- 家族、友人、パートナー、メンターと弟子のような、特別で個人的で親密な関係は、成立する前に「この相手は自分にとってどれだけ利益・損害をもたらすか」というような態度で臨んでいては構築できない。
- 信頼に基づく協力関係は、まず理由もなく、信頼することから始まる。人間は信頼されればそれに応えようと努力するものである。つまり信頼がきっかけになって、相手が本当に信頼に足る人間になるということがある。

リスク分析が失敗するケース

囚人のジレンマ

		プレイヤー2の選択	
		協力	裏切り
プレイヤー1の 選択	協力	(4, 4)	(0, 6)
	裏切り	(6, 0)	(2, 2)

括弧内の数値は、左がプレイヤー1の得点、右がプレイヤー2の得点を表す。

要点

- 人工知能は、人間を対象にした確率論的リスク分析のための革新的なツールである。
- 人工知能は現在、大金持ちを超大金持ちにすることに最も有効に活用されている。
- 一方、人工知能はしばしば社会的弱者をさらに不利な状況に陥れる。
- アルゴリズムには作成者の偏見や先入観や意見が、データには社会の持つそれらが反映される。
- 「人工知能だから公平」は大きな間違い。
- 人工知能は、他者を機械可読なデータと、そこから推論される属性の束とみなし、リスク分析の対象として扱う態度を助長する。