

横浜国立大学 IMS リスク共生社会創造センター
第14回シンポジウム

「リスク共生の視点からみた社会課題と対応」

AIと社会のリスク

2024年3月27日

横浜国立大学 総合学術高等研究院 リスク共生社会創造センター 特任教員(助教)
NPO法人 リスク共生社会推進センター 先進技術リスク検討WG リーダー

鈴木智也



リスク共生社会推進センター
Risk Symbiotic Society Co-Creation Center

0

本日の発表内容と本WGの意義

■ 目次

- 先進技術リスク検討WG(以下、本WG)で対象とする先進技術
- 「AIと社会のリスク」を考える上で重要な視点
- “従来技術”と“AI”との違いの整理
- 本WGでの当面の取組(AIを『機能』として捉えた整理)

■ 本WGの意義

今後の社会やあらゆる技術システムにおいてDX関連技術がインフラを形成する中で、
従来工学システムにおけるリスクの捉え方とは

大きく異なるアプローチを取らなければならない可能性がある

⇒ 先進技術である「DX」「AI」を題材に、リスクの捉え方や分析・評価手法の検討を開始

1

本WGで対象とする先進技術

■ AI: Artificial Intelligence (人工知能)

人間の知的な振る舞いをソフトウェアを使って人工的に再現したシステム

タイプ	機能	事例
特化型AI (ANI: Artificial Narrow Intelligence)	人間の知能の一部に特化した機能を持ち、特定の決まった作業を遂行するAI	<ul style="list-style-type: none"> 自動運転(Uber) 画像認識による推定や変換 文章生成(ChatGPT) 音声認識(Siri)
汎用型AI (AGI: Artificial General Intelligence)	人間の知能の総合的な領域の機能を持ち、人間のように感じ、考え、振るまう自己意識や精神を備えたAI	今はまだ存在しない (ChatGPT-4等の出現によって実現の可能性が見えてきた?)

※哲学者J.R. Searle(米国)が『Minds, brains, and programs』(1984)にて考案

■ DX: Digital X-formation (デジタルトランスフォーメーション)

「企業がビジネス環境の激しい変化に対応し、データとデジタル技術を活用して、顧客や社会のニーズを基に、製品やサービス、ビジネスモデルを変革するとともに、業務そのものや、組織、プロセス、企業文化・風土を変革し、競争上の優位性を確立すること。」

※経済産業省「デジタルガバナンス・コード2.0」より



Copilotにより生成
(人間の脳をシルエット
画像で描くよう指示)

2

本WGで対象とする先進技術

■ AI: Artificial Intelligence (人工知能)

人間の知的な振る舞いをソフトウェアを使って人工的に再現したシステム

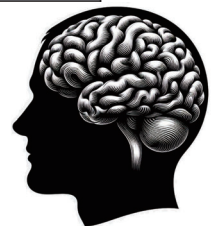
タイプ	機能	事例
特化型AI (ANI: Artificial Narrow Intelligence)	人間の知能の一部に特化した機能を持ち、特定の決まった作業を遂行するAI	<ul style="list-style-type: none"> 自動運転(Uber) 画像認識による推定や変換 文章生成(ChatGPT) 音声認識(Siri)
汎用型AI (AGI: Artificial General Intelligence)	人間の知能の総合的な領域の機能を持ち、人間のように感じ、考え、振るまう自己意識や精神を備えたAI	今はまだ存在しない (ChatGPT-4等の出現によって実現の可能性が見えてきた?)

※哲学者J.R. Searle(米国)が『Minds, brains, and programs』(1984)にて考案

■ DX: Digital X-formation (デジタルトランスフォーメーション)

「企業がビジネス環境の激しい変化に対応し、データとデジタル技術を活用して、顧客や社会のニーズを基に、製品やサービス、ビジネスモデルを変革するとともに、業務そのものや、組織、プロセス、企業文化・風土を変革し、競争上の優位性を確立すること。」

※経済産業省「デジタルガバナンス・コード2.0」より



Copilotにより生成
(人間の脳をシルエット
画像で描くよう指示)

3

「AIと社会のリスク」を考える上で重要な視点



Institute for
Multidisciplinary
Sciences
Yokohama National University



■これからの社会に必要な「リスクに対するスタンス」

「新しい技術が社会に導入される前にリスクを考えること」

「今後の社会像の変化を先取りしてリスクを考えること」

- ・「先んじて失敗して学習し、改善する」方法は大いに有効である
- ・しかしそれは、「致命的な失敗でない限り」という条件付き
- ⇒リスクと共に生きる社会を志向するならば、事前に対応できるなら、できることはやっておきたい。
最大の後悔を最小化したい。

⇒「“従来技術”と“AI”との違い」を整理

- (1)技術の成長・発展速度の違い
- (2)技術を提供する立場/利用する立場それぞれの目線の違い
- (3)技術を利活用する人間の性質の違い
- (4)影響の範囲・質の違い

4

“従来技術”と“AI”との違いの整理



Institute for
Multidisciplinary
Sciences
Yokohama National University



(1)技術の成長・発展速度の違い

従来技術に比較してAIは、

- ・技術の開発・成長速度が速い
- ・新しく社会に提供されるサービスがほとんど
- ・今後もその傾向は変わらないと予想される
- ・現時点で導入・利用実績や知見が蓄積していないものも多数

⇒リスク分析手法が未だ定まらず、リスクを考慮しないままに導入が進む傾向にある

⇒「対象システムを明確化した上でリスクを分析していくアプローチ」が、
通用しない、またはそれでは対応できない可能性がある

リスクの捉え方やリスク分析・評価の方法自体を研究対象とする必要がある

5

“従来技術”と“AI”との違いの整理

(2)技術を提供する立場/利用する立場それぞれの目線の違い

「AIを社会に提供する立場にとってのリスク」だけでなく、

「提供されるAIを利用する立場にとってのリスク」についても検討が必要

AIを社会に提供する立場 (AIサービス開発企業など)にとってのリスクの例	AIを利用する立場 (その他の企業・個人など)にとってのリスクの例
【AIを社会に提供しない場合】 <ul style="list-style-type: none">•競合他社に比較し、技術的な遅れや市場での競争力を失う•サービスや製品の質が低下し、顧客満足度が低下する•AI技術の進歩やイノベーションの機会を逃し、将来的な成長の機会を失ったり、長期的にAI技術の進展から取り残され業界内での立場が弱まる など	【AIを利用しない場合】 <ul style="list-style-type: none">•競合他社に比較し、効率化や生産性向上の機会を逃し、技術適応の遅れや市場での競争力を失う•情報分析能力を活用できず意思決定の質が低下する•AIにアクセスできない企業や個人が社会的・経済的に不利な立場に置かれる など
【AIを社会に提供する場合】 <ul style="list-style-type: none">•データの漏洩や不正使用、AIによる倫理的に問題のある判断や行動により、社会的な非難や法的な責任を負う•AIの誤動作やシステム不具合により、サービスの信頼性や品質が損なわれる•AI技術への社会的不信感や反発が強い場合、市場での受容が困難になる など	【AIを利用する場合】 <ul style="list-style-type: none">•AIが既存の作品や商標を含むコンテンツを生成し、知らずに著作権や商標権を侵害する•意思決定プロセスの説明責任を果たすのが難しい•人々の知的能力・判断能力がAIに依存することで衰退する•AIに過度に依存することで、システム障害時に業務に深刻な影響を及ぼす など

※各項目: Microsoft CopilotおよびChatGPTにより生成された回答を再整理

6

“従来技術”と“AI”との違いの整理

(3)技術を活用する人間の性質の違い

従来技術に比較してAIは、「誰にでも使える」ことがリスクになる可能性がある

- AIは、「誰がどのように使うか」によってリスクが大きく異なると考えられる
- 「わかった気になって」使ってしまう

従来技術では、

- 特定の条件を満たす人しか利用できないようにする
- 技術そのものに何らかの方法でリミッターをかける

などの方法をとってきた
しかし、AIに関してこれらのアプローチは可能だろうか？

AI技術単独ではなく、使用環境や利用する人間の性質の違いを考慮した分析が必要

(4)影響の範囲・質の違い

従来技術に比較してAIは、その影響が「物理的影響」とは異なる可能性がある

- 従来技術(工学システム)は、主に人間に対する物理的影響を対象にしてリスク分析されてきた
- しかし、AIの影響は物理的影響のみでは語れない

7

ここまでのまとめと検討が必要な論点

■「“従来技術”と“AI”との違い」のまとめ

- (1)技術の成長・発展速度の違い
- (2)技術を提供する立場/利用する立場それぞれの目線の違い
- (3)技術を活用する人間の性質の違い
- (4)影響の範囲・質の違い

リスクの捉え方やリスク分析・評価の方法自体を
研究対象とする必要がある

■上記に関して検討が必要な論点

以下について「なぜ考えるか」「何を考えるか」を整理する

- 何がリスクなのか？(リスクをどう捉えるか？)
- どのマネジメント主体の立場で見たときの、どの主体に対する影響のリスクについて考えるか？
- 実現している世界像(技術像・社会像)
- 上記世界像を想定するAI技術の開発・発展フェーズ
 - ・研究(研究シーズ) ・開発(試作) ・事業化(商品化) ・産業化(ヒット)

8

本WGでの当面の取組

■リスクの捉え方の案:「AIを『機能』として捉える」

※単に「どんな機能があるか」ではなく、

「どんな機能を持たせようとしているか」「今後どんな機能を持ちうるか」という方向性・志向性で捉える

○「AIを『機能』として捉える」をなぜ考える？

- 必ずしも、「対象システムを明確にしないとリスクが分析できない」わけではない
- 「機能」として捉えれば、対象システムそのものを詳細に知らずとも一定程度の分析が可能になり得る

○「AIを『機能』として捉える」で具体的に何を考える？

- AIは、今後どんな機能を持ちうる技術だろうか？
 - ・何らかの形で人間の持つ能力を代替・拡張する機能を持つ
 - ・従来の「身体的機能」を代替・拡張する技術とは大きく異なることが特徴
- 「AIを『機能』として捉える」捉え方によって、分析できること/できないことなどを整理

9

AIの『機能』についての整理(途中経過)



Institute for
Multidisciplinary
Sciences
Yokohama National University



■現状の整理状況

•AIが持つ/今後持ちうる「機能」

•人間の頭脳の機能を代替する

•例①:「入力されたデータ・情報を、何らかの評価軸に従って評価・判断する機能」

•人間の頭脳の機能を拡張する

•例②:「人間の記憶や認識を補完・拡張する機能」

•XXXXXXXX?

•例③:「XXXXXXXXXXXX」 …… (これらについて、今後体系化していきたい)

10

まとめと今後の展開



Institute for
Multidisciplinary
Sciences
Yokohama National University



■まとめ

•DXやAIを題材に、新しい技術が社会に導入される前に

リスクを分析する方法論について検討を始めた

•「従来技術」と「AI」との違いについてまとめ、

AIを『機能』として捉えた上で、その機能について整理し始めた

■今後の展開

•AIの『機能』を体系的に整理し、この捉え方の妥当性を議論

•その他の各論点も考慮し、AIのリスクについて整理

DX・AIなどの先進技術リスクについてご興味ある方、一緒に考えてくださる方
ぜひご参画をお待ちしております

11