

ムーンショット目標3

「2050 年までに、AI とロボットの共進化により、自ら学習・行動し

人と共生するロボットを実現」

目標3 ELSI 論点整理

令和 7年 11 月

目標3プログラムディレクター:福田 敏男

目次

(1) 本資料の目的.....	3
(2) 本資料の対象範囲	3
(3) 本資料の対象読者	3
(4) 目標3の各プロジェクトの ELSI 課題の概要.....	4
(5) 本資料の構成.....	5
1. AI とロボットをめぐる ELSI についての考え方の前提	6
(1) 目標3が想定する AI ロボットの範囲と ELSI.....	6
(2) AI とロボットを個別に考えるのか、それとも AI ロボット(一体化したもの)として考えるのか	6
2. AI とロボットの自律性の進展の度合い(2030 年 2050 年)と事故・事件が起こった場合の責任の主体について	8
(1) 自律性の進展の度合いの想定(仮説)	8
(2) 進展の度合いに伴う ELSI 上の問題点は何か	9
(3) ELSI の問題点のうち責任の分界点はどこになるか	14
3. AI ロボットに関するデータガバナンスの ELSI 上の問題点について.....	16
4. 月面での活動における ELSI 上の問題点は何か.....	19
5. 社会的な受容性の観点	26
参考文献.....	31
付録	32
付録1. 目標3 AI ロボットの ELSI 論点整理の課題の俯瞰図	32
ELSI 研究会関係者	33

はじめに

(1) 本資料の目的

本資料は、ムーンショット目標3「2050 年までに、AI とロボットの共進化により、自ら学習・行動し人と共生するロボットを実現」における倫理的・法的・社会的課題(Ethical, Legal and Social Issues:ELSI)の論点を整理することが目的である。論点の整理においては、目標3の各プロジェクトの ELSI に関する共通項を見出して整理することに加え、2050 年及び 2030 年の時点で予想される ELSI を論じることを目指した。

なお本資料の目的は論点を整理するまでであり、細かなガイドラインを示すことまでは目的としていない。

本資料は、目標3の中に設置した ELSI 研究会の議論を踏まえて作成している。目標3の ELSI 研究会とは、目標3の友枝アドバイザーが統括となり、複数の外部有識者にも参加していただき、各プロジェクトとの議論や外部有識者同士の議論を行う会議体である。

(2) 本資料の対象範囲

ムーンショット目標3の各プロジェクトの AI ロボットに関する ELSI を対象とし、令和 6 年度及び令和 7 年度上期に開催した「目標3 ELSI 研究会」において各プロジェクトと議論した内容を基に論点を整理している。

なお現在の8つの研究プロジェクトを議論の土台とするが、将来に研究される可能性のある AI ロボットも可能な範囲で想定対象としている。

(3) 本資料の対象読者

本資料は、主に AI ロボットの ELSI に関心のある研究者を対象読者とするが、研究者以外の一般人にも関心を持って読んでいただければ幸いである。ただし、AI ロボット関連の知識を一定程度有する研究者を基本読者としているため、AI やロボットの専門的な説明が十分に行き届かない箇所がある点をご容赦いただきたい。本資料では、ELSI 論点整理を主眼としているため、目標3プログラム及び各プロジェクトに関する説明は必要最小限に留めている。詳細は JST のムーンショット目標3の紹介ページを参照いただきたい。

(4) 目標3の各プロジェクトの ELSI 課題の概要

目標3プログラムでは、3つのターゲット領域に対して、それぞれ2～3個のプロジェクト、合計8つのプロジェクトがある。それぞれのプロジェクト名称・掲載ホームページ・ELSI 課題例を示す。8つのプロジェクトそれぞれの AI ロボットの種類は様々であり、人型スマートロボット、複数の形態を提供する AI ロボット群、科学実験を行うマニピュレータ型、災害現場の建機型、月面拠点構築の各種ロボット群などがある。詳細は掲載ホームページ(HP)をご覧ください。

ターゲット	プロジェクト	各プロジェクト名称・掲載HP・ELSI課題例
<ターゲット1> 人に寄り添うAI ロボット	菅野プロジェクト	一人に一台一生寄り添うスマートロボット https://www.jst.go.jp/moonshot/program/goal3/31_sugano.html 一台のスマートロボットによる医療・介護利用でのELSI課題
	平田プロジェクト	活力ある社会を創る適応自在AIロボット群 https://www.jst.go.jp/moonshot/program/goal3/34_hirata.html 複数の適応自在ロボットによる医療・介護利用でのELSI課題
	下田プロジェクト	主体的な行動変容を促すAwareness AIロボットシステム開発 https://www.jst.go.jp/moonshot/program/goal3/39_shimoda.html Awareness AIロボットによる医療・介護利用でのELSI課題
<ターゲット2> 科学探究を行う AIロボット	原田プロジェクト	人とAIロボットの創造的共進化によるサイエンス開拓 https://www.jst.go.jp/moonshot/program/goal3/33_harada.html AIロボット科学者によるサイエンス実験でのELSI課題
	牛久プロジェクト	人と融和して知の創造・越境をするAIロボット https://www.jst.go.jp/moonshot/program/goal3/36_ushiku.html AIによる学術論文作成・論文査読・論文著作権のELSI問題
<ターゲット3> 難環境で活動 するAIロボット	永谷プロジェクト	多様な環境に適応しインフラ構築を革新する協働AIロボット https://www.jst.go.jp/moonshot/program/goal3/32_nagatani.html 災害現場での協働AIロボットによる被害者救済をめぐるELSI問題
	國井プロジェクト	未知未踏領域における拠点建築のための集団共有知能をもつ進化型ロボット群 https://www.jst.go.jp/moonshot/program/goal3/38_kunii.html 月面での進化型ロボット群による活動におけるELSI課題
	吉田プロジェクト	月面探査／拠点構築のための自己再生型AIロボット https://www.jst.go.jp/moonshot/program/goal3/3B_yoshida.html 月面での自己再生型AIロボットによる活動におけるELSI課題

(5) 本資料の構成

「目標3 ELSI 研究会」の議論を踏まえて設定した5つのテーマ(1章～5 章)について論点を整理した。また各テーマの章においては、以下の項目に分けて情報を整理した。

課題の表題には、【プロジェクト共通】、【菅野プロジェクト】、【吉田プロジェクト、國井プロジェクト】といった記載により、その課題が共通課題なのか、プロジェクト固有の課題なのかを示している。

また、必要に応じて、【用語説明】、【注釈】による説明を加えている。

【関連情報・キーワード等】

【論点の整理】

<倫理的課題>

<法的課題>

<社会的課題>

【本章のまとめ】

1. AI とロボットをめぐる ELSI についての考え方の前提

【関連情報・キーワード等】

- ELSI：倫理的・法的・社会的課題（Ethical, Legal and Social Issues）
- AI：人工知能（Artificial Intelligence）

【論点の整理】

(1) 目標3が想定する AI ロボットの範囲と ELSI

ムーンショット目標3において研究している AI とロボットに関する ELSI を対象とする。現在の8つのプロジェクトの AI とロボットを議論の前提としつつ、それ以外にも目標3の想定範囲内と思われる AI とロボットも議論の対象とする。

本資料では、各プロジェクトの AI とロボットについての具体的な説明は省略するが、概ね以下の形状のロボットが研究開発されている。

ターゲット1：人と共生して活動する自律移動型ロボット（ヒューマノイドロボット）、医療（診断を含む）で使用されるロボット、介護支援などのために人に装着して動作するロボット。

ターゲット2：科学研究の実験等を行うためのマニピュレータ型のロボット（単体あるいは複数マニピュレータ）。

ターゲット3：難環境で活動する自律移動型ロボット。月面探査用の小型の群ロボット、月面活動用のマニピュレータの形状組み替えが可能な移動ロボット、地上災害対策用の自動建機ロボットなど。

(2) AI とロボットを個別に考えるのか、それとも AI ロボット（一体化したもの）として考えるのか

<背景>

AI やロボットに関する ELSI については、すでに様々な議論がなされている。一例として、国立研究開発法人科学技術振興機構（JST）研究開発戦略センター（CRDS）が 2023 年 5 月に公表した「科学技術・イノベーションの土壌づくりとしての ELSI/RRI 戦略的な科学技術ガバナンスの実現に向けて¹⁾」の資料がある。その「付録 C ケーススタディ」の8項目の中に、「(1)人工知能(AI)、(4)サービスロボット、(6)医療・ヘルスデータ」があり、想定される ELSI 課題の例を記載しているが、AI とロボットを分けて論じている。また最近では、2024 年の欧州 AI 法 (AI Act)²⁾をはじめ、特に AI に関する議論が各国で進んでいる。日本政府では、2022 年 4 月に策定された「AI 戦略 2022」を起点に、AI に関する政策と制度の整備を段階的に進めている。（日本政府の動きについては、第 2 章末尾の【注釈】に記載しているのでご参照いただきたい。）

このように AI 単独の議論が世の中で進んでいる中で、目標3の AI ロボットにおける ELSI の論点を AI 単独とロボット単独で議論・整理するのか、あるいは AI とロボットを一体化して議論するのかを整理する。

<本論>

AI とロボットは、それぞれに発展の歴史がある。AI はコンピュータ上の仮想空間で動作するソフトウェア（プログラム、データベース等）であった。一方、ロボットは現実世界の環境で動作する物理的な身体を持ち、その動作は何かしらの制御により行われるものであった。この制御の仕組みとして制御用ソフトウェアが使われてきたが最

近では AI の活用が進展してきた。なお、仮想空間上のロボットも想定されているが、ここでは一旦議論の枠外におく。

AI は、第 3 次 AI ブームにおいて機械学習(マシンラーニング)や深層学習(ディープラーニング)が盛んになり多くの研究開発が行われてきた。その後、2022 年 11 月に米国の OpenAI 社が公開した「ChatGPT」が世界的に注目を集めて以降は第 4 次 AI ブームに入ったとも言われており、大規模言語モデル (large language model、LLM) や生成 AI (generative artificial intelligence) などの研究開発が盛んに行われている。こうした AI の発展に伴い、AI そのものに関する ELSI の議論が欧米や日本でも活発に行われている。

一方、従来のロボットは、動力学などに基づいたロボット制御が多く使われてきたと認識されており、ロボットを制御するためのソフトウェア(プログラム)はロボット研究分野の中で独自に開発されてきたと言える。しかしながら最近では AI の発展に伴い、ロボットの動作に AI を活用する研究開発が多くなっている。最近開催されるロボットに関する国際学会(ICRA、IROS)においても、ロボットの制御に各種の AI を適用した研究が多く見受けられている。最近のロボット研究では AI を抜きにしては語れない状況となっていると言える。また現実目標3のプロジェクトにおいても各種の AI を適用したロボットが開発されている。

<参考:国際学会の名称>

•ICRA: The IEEE International Conference on Robotics and Automation

•IROS: The IEEE/RSJ International Conference on Intelligent Robots and Systems

このように従来は AI とロボットはそれぞれ異なる発展をしてきたが、ここ数年ではロボットの制御に AI が欠かせないものになっていると言えよう。もし、ChatGPT のような知識を扱う AI と、ロボットの制御の AI を分けて考えるのであれば、AI の ELSI とロボットの ELSI をそれぞれ別に議論することになるであろう。また AI の範囲をどう捉えるかの議論もあるが、ロボットを自動で動かすところを AI と見なせば、動くもの(ロボット)と AI を一緒に考えないといけないことになる。さらに最近では、現実世界の環境を認識して何らかの物理動作を行わせるような AI 技術も進展しており、厳密な定義を省略すれば、それらは Physical AI や Embodied AI といった用語で呼ばれている。なお目標3においては、プログラム開始当初から「Embodiment AI・Robot の概念」を提唱している。そのため目標3では、何らかの AI を活用してロボットが動くことが多いため、AI とロボットを一体として合わせて議論するのがよいと考える。

なお一体化して議論する場合にも、分けて考える場合に比べて何かを見落とす懸念はないものと考えている。ロボット単体に関しては、例えばロボット工業会などでいろいろな基準があり、法律的にもアクチュエータの仕様も規定があるため、目標3でこれらのロボット単体の議論をするのは得策ではないと考える。また AI 単体の議論(欧州の AI Act など)は今後も欧米や日本で議論が継続されていくが、目標3の AI ロボットに関連する AI 事項は必要に応じて確認していくことになる。

【本章のまとめ】

- 目標3においては、AI とロボットは個別に議論せず、一体化した AI ロボットとして ELSI を考える。

2. AI とロボットの自律性の進展の度合い(2030 年 2050 年)と事故・事件が起こった場合の責任の主体について

【関連情報・キーワード等】

- 自律型ロボット
- 製造物責任法
- 知的財産基本法、特許権（特許法）、著作権（著作権法）
- 欧州（EU）：AI 法「Artificial Intelligence Act (AI Act)」

【論点の整理】

(1) 自律性の進展の度合いの想定(仮説)

目標3の8つのプロジェクトにおいて、2030 年から 2050 年までの段階的な進展をそれぞれ計画しているが、概ね要約すれば、2030 年が条件付き自律化 AI ロボット、2050 年が完全自律化 AI ロボットといった想定である。

自律型ロボットと呼ばれるものは過去から現在において複数存在しており、2024 年現在では産業用ロボットにおいても決まった動作の繰り返しだけではなく、状況変化が生じた場合に人間が逐次指示をしなくても、ある程度の範囲で自律的に動作するロボットも開発されている。目標3が目指すものは、現在の自律型ロボットと呼ばれているものよりも、さらに自律機能を向上させたものである。

2030 年では一部の遠隔操作など何かしら外部から支援を行いながら、かなりの部分を自律的に動作するようになるだろう。

2050 年では目的を与えると AI ロボットが目的達成の手段として、どういう経路を辿って到達するかは、AI ロボットのシステム側に自由度があり自律的に動くだろう。ただ 2050 年の完全自律化においても、AI ロボットが自分で目的そのものを設定することが果たして本当に実現するのかどうかは分からない。

なお、ここでの AI ロボットの自律性は、哲学の分野において(例えばカントのような哲学者によって)論じられている「自律性」の概念とは異なっている。AI ロボットには、「自律性」についての哲学分野での議論で重視される「意識・理性・道徳など」が備わっていないためである。2050 年においても人間と同等レベルのものが AI ロボットに備わるとは想定しづらいため、AI ロボットの自律性は限られた範囲での議論となる。ここでの議論は、あくまでもロボット工学分野における自律性の議論に留めておく。

想定される 2030 年の AI ロボットと 2050 年の AI ロボットの自律性の違いについて少し付言する。2030 年の AI ロボットはあくまでも限定された文脈、場所、時間、選択肢の中で自律的に(操作者による逐一の介入なしに)動作するが、一方 2050 年の AI ロボットはオープンエンドな環境で、あらかじめ限定されていない選択肢の中から意思決定をして行動するかもしれない。前者はいわば囲いの中に閉じ込められた状態、後者は野に放たれた状態で自律的に動作している。そうなれば後者の自律的 AI ロボットに対してはより高度な柔軟性、安定性、信頼性が求められると同時に、想定外に備えた安全機構、補償、保険、契約、コンセンサスなどが必要となるだろう。

(2) 進展の度合いに伴う ELSI 上の問題点は何か

AI ロボットの自律性の進展に伴う ELSI 上の主要な問題点は、事故発生時の責任の所在の観点である。一部の遠隔操作など、AI ロボットが、あくまで道具であり人間がコントロールしているという前提に立てば今までの法律の解釈で操作する人間側の責任ということで対応できるだろう。一方、AI ロボットが完全に自律的になり、初期値と違う行動基準を持つということになってくると、AI ロボットが何か問題を起こした時の責任について考える必要がある。参考までに自動運転の車の場合は事故を起こすと、自動運転を作った会社、実行した自治体、AI のプログラマー、などが悪いといった責任論となる。このような責任について AI ロボットの基本的な考え方をまとめたい。

2030 年の技術レベルと 2050 年代とは違うはずなので、区分して考えたい。古い話で言うと、ラジオの時代、テレビの時代、白黒テレビの時代、カラーテレビの時代があり、それに応じて人間のテレビ視聴の仕方は変わってきている。車もかなり進歩してきている。これらの歴史的事例を念頭におきながら、社会状況と関連付けて考えたい。

例えば現時点では、学習するためのデータセットは、基本的に人間が用意して与えている。今後ロボットが、あちこち歩きながらデータを収集して、その収集したデータに基づいて学習し、その学習によってロボットの行動を変化させ、それで変化した周囲のデータを収集するというようなことをやっていくようになれば、かなり自律性は高まっていくだろう。イギリスの情報倫理の専門家であるフロリディは、機械は自律性を持ったり、行為者性つまりエージェンシー(agency)を持ったりすることは可能だが、ロボットのレスポンシビリティ(これから起こることへの責任)については別であり、必ずしも考慮しなくてもよいという見解を示している。

今までの法律体系はおそらく行為者の行為・行動とレスポンシビリティあるいはアカウントビリティ(すでに起きたことへの説明責任)を結びつけていたと思う。レスポンシビリティをどう考えるかということについて議論をしていく必要がある。

例えば今の手術ロボットは人の手の動きを再現するに留めていて、医者の手動きを再現するだけなので、ロボットが何かしたとしても責任は医師にあるとなっている。一番クリティカルであるはずの手術ロボット分野でもまだ議論が進んでいる途中であり、どちらの責任かという話まで発展していない。そのため手術ロボットにおける自律化したロボットの責任問題の議論は今後の参考になるかもしれない。

例えば宇宙で科学実験をしたいというような例を考えると、人間の判断は重力下でのカン・コツに基づいている。宇宙の遠いところだと人間の判断を伝えるにも時間がかかるので、そういう極限的な環境においては、そのロボットが微小重力下で動きやすいロボットの動きで考えた実験をすることは十分ありうると思う。ロボットの自動と自律を使い分けることが多いが、自動は人が判断したものをロボットが実行するもので、自律はロボット自身が判断するというものである。目標3の AI ロボットは、細かい判断であれ大きな判断であれ、自ら判断していくことを目指していると考えている。

文化による考え方の違いがあると思うが、特に西欧キリスト教型の文化は(AI ロボットの)自律を嫌うところがある。戦争法を考えている時にも、あくまでも人間が判断をしなくてはいけないという考えがまだまだ強いところがある。このように ELSI の問題も国や文化による違いがあると予想されるが、本資料では、まず日本における問題点に焦点をあてることにしたい。

<倫理的課題>

事故発生時の責任の所在【プロジェクト共通】

AI ロボットの自律性が高度化して、どのような社会の場面で使われるかは今後の 2030 年～2050 年における社会への実装の進展による。民間企業がビジネスとして AI ロボットを社会の中で提供していく場合において、どのような利用シーンが想定されるかを完全に予想することは難しいが、目的を設定するのは人間であるため、その動作による事故発生時の責任も人間および関連する組織に起因することが考えられる。

ただし、作業目的そのものを自ら生成して考えていくような AI ロボットが将来に存在するかもしれない。そうすると、イーロンマスクなどが盛んに言及している、いわゆるエグジステンシャルリスク(人類の存亡にかかわるリスク)が生じる可能性があり、そうすると、それに対応すべき ELSI を視野に入れる必要があるだろう。

<法的課題>

AI ロボットの人格の問題【プロジェクト共通】

AI の進展により AI ロボットに人格のようなものを認めるのか否かという議論がある。現時点では、**法人格の問題を真剣に取り上げている論者は少ない**。AI やロボットに法人格を与えると、その背後に存在する人間の悪意がある場合でも責任を追及できなくなるというのが主たる理由である。ただ将来の状況変化によっては、議論が繰り返される可能性はある。とはいえ AI ロボットのうち、月面探査用の小型ロボットの法人格を議論することはないだろう。人型のヒューマノイド AI ロボットが人間に近い存在になった場合に議論が再燃される可能性はあるかもしれない。

AI ロボットの法的な規制【プロジェクト共通】

この問題の前提として、現在の法規制が人間を対象として設計されていることを踏まえる必要がある。人間は規制を(本来の意味で)適当に解釈して行動するので、例えば小学生が正月にお屠蘇を飲んだからといって警察官に咎め立てられるようなことはない。このように現場の公務員により弾力的な法運用がなされることを、行政学では street level bureaucracy と呼ぶ。だが AI ロボットはプログラムを介して遵守するよう求められた規則を忠実に適用するだろう。その結果、人間が運転する周囲の自動車がすべて制限速度をある程度超過して運行しているような状況でも自動運転車だけが制限速度を守り、渋滞や衝突の危険を生じさせるといった問題が起きることが予測される。

これを回避する方法は 3 通り考えられる。第一に、AI ロボットにも人間と同様に規則を適当に解釈し弾力的に対応する能力を付与すること。しかしこれは AI ロボットの動作がプログラムによる事前指示に忠実ではなくなることで、その結果が予測困難になることを意味するだろう。第二に、法規制を AI ロボットが忠実に適用しても問題ないレベルまで精密なものへと修正していくこと(例えば二十歳未満の者の飲酒の禁止に関する法律に儀式や年中行事に関する例外を書き加える)。だがこれには AI ロボットの関与しない分野も含めて膨大な法改正のための努力が求められることになるだろう。第三に、人間のための法規制とは異なるものとして AI ロボットに向けた規制を位置付け、両者の優先順位などを明確にすること。例えば自動運転車が常に制限速度を守り一定のルートを実行することを前提として、そこから生じる問題に弾力的に対処する能力を持つ人間にそれに対して配慮する義務を課す(自動運転車に対する回避義務、運行ルート上の駐車禁止など)ということが考えられる。

いずれにせよ、人間という概ね単一の特性を持つ個体群を前提として組み立てられてきたこれまでの法規制のあり方に対しては根本的な再検討が必要になるものと考えられる。

自動運転の場合は、社会の中で使われる前提で、自律性が高まっていく度合いによって、事故が起きた場合の責任・保障・刑罰の問題で分けられていると思う。どういう状況でどういう用途で AI ロボットを使うかという観点でレベル分けを考えることが必要である。例えば EU の AI Act ではリスクを分けており、許容できないリスクからマイナーなリスクまでレベル分けをして、そこに用途と機能が結び合った表現になっている。つまり、こういう用途で使う、こういう機能は非常にハイリスクであるというような表現になっている。

一般的に法的な観点から、誰に責任を負わせるか、誰が管理するかといった点で、想定される自動建機(AI ロボット)の自動化のレベルによって責任者が変わる。複数のドローンを飛ばすときに、どのくらいの人に操作させるか、閾値をどう決めるか、システムの仕様によって変化する。自動化や自律化の範囲を限定するのであれば、今の枠組みで整理できるだろうが、ムーンショットで現在は夢みたいなものであっても将来実現されるものまで考え、産業構造が大きく変化した場合、規制の考え方も変わってくるのが予想される。

流動性が高い状況下だと、ソフトローでフレキシブルに考える場合もある。ハードローは 10 年先まで影響を与えてしまうため、どういう風に設計するのがよいか、技術がどこまで進展するか不明な段階では、ソフトローに依拠するアプローチが有効になる場合がある。ただし、人権侵害の禁止など、明らかにハードローで規制する必要がある面もあるため、ソフトローとハードローを上手く使い分ける必要がある。なお EU は特有の哲学に基づいて出口を予め決めており、ここには、発想やアプローチの違いがある。

また保険賠償と製造物責任は、多くの国で実現している。保険事故が生じた際に、被保険者が作ったモノが本来の性能を備えていない場合には、保険会社が支払った保険金を被保険者に求償できるので被保険者がモラル・ハザードを起こさないという仕組みが、保険が適切に機能する前提である。しかし、全世界で問題提起されているのは、保険会社が最先端の AI を理解できず、事実上十分に被保険者に求償ができないという構造が存在することである。結果として保険が機能しないのではないかという課題がある。いずれにしろ、知的財産権と産業構造の問題が関係してくるので、法制度を構造化して保険がうまく機能するように進めることが必要である。

ちなみに生命保険の場合は、がん・怪我・病気になったら、掛け金を払っていれば、(何らかの条件はあるが)保険金が支払われる。例えば、生命保険のイメージで考えると、ロボットと共生する社会では、ロボットが何かしたこと、あるいはロボットが人間に危害を加えた場合に、その保険の中から自動的に支払われるとする。それに対するリスクを個人が負担する、もしくは、どこかが負担するのかというような議論が発生する可能性がある。自動運転も誰かが保険を担うと思われるが、過失割合が明確に定義できないことは問題である。

次に、ロボットと人とのインタラクションについて、例えば菅野プロジェクトの人型ロボットでは、2030 年になると接客や拭き掃除や一部介護などで、人とのインタラクションがあるが侵襲性はないという想定であるが、2050 年になると、個々人の執事として働くことになり、そこでのインタラクションには侵襲性があると想定される。

ロボットが人とのインタラクションを伴う場合に、人間側の感情を無視するのは難しく、AI Act の 5 条の f) に何らかの形で該当してしまう懸念がある。インタラクションが、結果的に誰かの判断や行動をゆがめて望ましくない結果をもたらした場合、5 条の a) あるいは b) がどういう意味を持ってくるのかは、今後さらなる調査が必要と思われるほど大きなテーマになりうる。AI Act では絶対禁止事項として 5 条 1 項で a から h まで 8 項目ある。a) のサブリミナルは、意図的に欺瞞的な行為で実質的に集団や人の行動をゆがめて重大な損害を与える、または恐れのある方法で意思決定を行わせることが禁止となっている。当初の a) 項は、自殺に追い込むサブリミナルの記載があり絶

対禁止だったが、議論が広がり、記載に「恐れ」が入っているので、広い範囲かもしれない。もう1つの④は職場および教育機関の分野において、自然人の感情を推測するために使ってはいけないとあり、これは相当広い。例えば職場で人事部門がアンケートして、機嫌の悪い人を探して声をかける場合があるが、こういったことも全部禁じられるかもしれない。そうすると例えば AI ロボットが会社で人とインタラクションして機嫌の悪い人を探してアドバイスするということはおそらく禁止されるであろう。我々のプロジェクト以外でも該当しそうな事項であり、なぜ、ここまでの内容が AI Act として記載されているかは、情報収集した方がよいであろう。もし欧州の AI Act の根底に、感情的な交流を伴うインタラクション自体に否定的な態度があるとする、そもそもロボットによるケアという方向性自体と相性が悪い。ロボットと人のインタラクションは目標3において重要な観点なので AI Act 及び今後日本で制定される同様の AI 規定には注意すべきである。

【注釈】なお AI ロボットの法人格については、責任問題の構造を簡素化するための道具として導入され得る点には注意する必要がある。人間が運転する自動車であっても事故発生時の責任の所在が状況によっては複雑な問題となり得るところ(例えば運転者が製造上の欠陥の存在を主張する場合)、その結論が出るまで被害者が十分な救済を受けられないような事態が発生することを避けるため、運行供用者の無過失責任が規定されている(自動車損害賠償保障法 3 条 1 項)。これは運行供用者(一般的な自家用車の場合にはその所有者)を第一義的な責任者として指定し被害者救済にあたせた上で、製造上の欠陥など他の理由が想定される場合には運行供用者からその担い手と目される主体に対する求償を行なわせることで、被害者の迅速な救済と責任分担の確定を両立させることを狙ったものと理解される。運行供用者が明確でない形で自動運転車が社会に登場した場合(例えば誰でも利用可能な公共交通として導入される)、自動運転車自体を法人として扱うことで運行供用者と同様の責任を負うように定めることは想定できるだろう。ここで問題になっているのはある存在を法的な権利義務の主体として扱うことにより社会的な便宜が生じるかという道具的人格性の問題であり、我々の社会の一員として相互承認の対象になる内実(理性・感情・受苦能力など)を備えているかという本質的人格性の問題ではない。

参照、大屋雄裕「AGI と社会と人類と」情報通信政策研究 8 巻 1 号(2024 年)³⁾。

AI ロボットが関連する知的財産権(論文ほか)の課題【牛久プロジェクト】

論文について、2030 年～2050 年に AI 技術が進んだ場合、論文は人間が書くのか AI が書くのか、また査読はどうなるのかという点について、牛久 PM は両方とも人間が行うものと考えている。AI が行うのは提示である。例えば、論文を書くときに、こういう法則に基づいて発見を行い論文としてこうなったという提示までは AI が行うが、それを論文投稿する意思決定は人間であると思っている。また査読する側としても、例えば論文の中でこういう実験が足りてないとか、こういう別の論文と重複しているということは、あくまで AI 側は提示を行うことまでで、良い点も悪い点も加味して論文誌や国際会議で論文をアクセプトするかどうかは人間が意思決定するものと考えてるのが現時点では適切である。この場合には従来通りのやり方なので ELSI の観点で言うと非常に考えやすい。

一方、著作権についていえばロボットが権利、例えば特許を得たとして、その特許収入を得たいのかというと、ロボットは別に特許収入を得たいわけではないかもしれない。もし AI ロボットが著作権とか、特許に対する権利とか、知財的な権利を持つとしても、人間がその権利を利用するのはだいぶ違う意味合いになると思う。その点をどのように考えていくのかという問題がある。人間の場合と権利も責任も全然違うものになるので、人間の方に分散したり分配したり、その社会制度の設計の問題が大きいので考えていかなければいけない。

著作権や知財を AI に帰属させないという考え方の前提は研究者をサポートする程度の AI しか用いないという状況であろう。将来、本当に自律的な AI ができて大量に学習させた状態で、ある研究者がたまたま使ったら確率

の問題で何らかの発明が出てくるかもしれない。その場合に、従来とは異なる知財の考え方も出てくるのではないだろうか。現実にはまだそこまで行かないと思われる理由は 2 点ある。1 点目は、人間の研究者がいなくなるのか残り続けるのかという点で、残り続けると想定している。もう1つは AI に知識をどんどん与えていくと発明が生まれるのは、ある意味その通りである。ただし、今後の特許の考え方次第だが、実施例を伴ってはじめて有効な特許になると考えるのがよいのではないだろうか。もし AI が着想したと思っても現実世界で実施しないといけない。合成の結果が良い薬になるかどうかの生体実験検証で時間がかかる。現実世界に検証できる基盤の設備を作らないと、AI の着想を知的財産権へとつなげる体制にはならない。

<社会的課題>

デュアルユースの問題【プロジェクト共通】

当たり前であるが、本人の意図と違ったところで技術が転用されることは、人類史上必ず起こってきたことである。このデュアルユースの問題をどう扱うか、悪用をゼロにはできない。最小限にする努力が必要である。技術が出来てからではなく ELSI の問題は一緒に議論が必要である。

デュアルユースの点、どういうときに国家の責任が問われるか、有害な干渉を他国に与えたかどうか、という意図が重視されている。AI の進展によって意図を何も国家が行わなかった場合など、EU の法制化がハードローで進んでいるので注視すべきである。両用技術製品は貿易問題としても大きな問題点だと思う。

責任に関する社会的な受容性の問題【プロジェクト共通】

自動運転の場合に誰に責任を負わせるかについては、法的又は倫理的検討を要するだけでなく、より広く社会的なインパクトに関する想像力を要する。例えば、人々の感情は重要である。被害者の納得も重要な論点であり、社会的な議論と話し合いが必要である。全体としての被害は減ったが、自動運転でなかったらこの人は死ななかったかもしれないというような場合でも、人々の生や感じ方をどのように考えるのかについて、議論が深められる必要がある。技術利用や制度や実践がひとたび慣習となればあえて問い直されない問題なのかもしれないが、広範な社会的インパクトに対する私たちの想像力と対応力とそれに応じた対応とが必要である。

将来、AI ロボットの自律性の度合いが高まるにつれて、責任に関する社会的な受容性の議論は変わってくるかもしれない。現在のように AI ロボットの開発企業やサービス提供企業の責任の議論になるのか、果たして AI ロボット自身への責任を追及したいという人間の感情が起こりうるのかなどについても、理論的な検討と経験的な調査の両輪が必要となる。

(3) ELSI の問題点のうち責任の分界点はどこになるか

想定外の事故: 自動運転による事故、診断ミス等

想定外の事件: 契約上のミス等

【注釈】言うまでもないが、自律的でなければ、責任はすべて関与した人間(設計担当者、プログラマー、AI やロボットの操作者等)になる。

これまではロボットへの動作指示は人間であり、力の制限や速度の制限の安全基準が設けられていた。しかし、動作を起こす決定そのものを人間ではなく、AI が行うという点を新しく考えていかなければいけない。AI が何らかのロジックで判断してロボットが危険な動きをする場合に、どのように備えていくのか、動作コマンドを生成するというタスクを人間だけでなく AI も担うところがこれからの一番大きな課題の一つである。

この問題はロボット固有の話でなく、AGI(汎用人工知能)や ASI(人工超知能)という AI の問題でもある。基本的には AI 分野の議論が、ロボットでも踏襲されると思われる。現在のロボットの AI が強化された時に何が変わるかという、基本的にシステムの自由度が上がることになる。また、AI ロボットが究極的に、工学的な意味での自律性を超えて、自ら目標を定めてそれに向かって行動をするという意味での自律性を獲得するのかについては慎重に考える必要がある。AI ロボットの行動については、基本的に誰か人が最初に目標を設定するはずである。その目標の実現に向けて行動する際のロボットや AI の自由度(裁量の範囲)は、今後どんどん上がっていくだろう。その目標の与え方と自由度の設定を誤って、自由度の中で AI ロボットが悪さを行うとしても、その目標と自由度を与えたのは人である。それを AI 自身が自己生成するようなことが本当に実現するかどうかは予見が難しいが、究極的にはそのような意味での自律性を AI が持つことはないのではないだろうか。とはいえ自由度が上がっていく中で、AI ロボットが自分で自分の機能を作り変えとか、自分でプログラミングしていくとか、いろいろな材料を使って自分で体を変えていくとか、やれることが変わってくる。その結果として与えられた目標に対する実現手段の中で、人に危害を及ぼすようなことが起こるかもしれない。自由度をどのように制約するかというのが、最終的には議論になるであろう。

AI ロボットが進化したとしても災害対応では、実際には初めての状況が発生するケースが多いので、人間が関与しないことはあり得ない。現時点は、人が外部から遠隔操作することが多い状況である。ただし取得した災害状況の画像情報から状況を判断する部分は AI を活用する。ロボットが自分で判断する状況はまだ先の話であり、10 年後に遠隔操縦(離れた場所から人が画像を見て操縦)で災害対応ができる状況になっていたら万々歳の認識である。阿蘇の大橋の復旧工事では遠隔操縦の機械が導入されているが、まだ他のケースでは難しい。10 年後に、自動で動く部分と遠隔操縦の中間くらい入っていればよい。

上記の議論等から、AI ロボットの適用領域によって、人間と AI ロボットの役割分担が変わってくるとされる。何かの目的に従って自律的に動作を行い、その結果、何らかの事故が生じた場合に、責任の分解点は、それぞれの役割に応じて関与した人間及び組織に割り振られるものと想像する。

【用語説明】

ソフトロー: 権力による強制力は持たないが多くの人がしたがっている規範。

公的なガイドライン(世界人権宣言など)、業界団体の自主ルール、企業の自主ルールなどのこと。

ハードロー: 権力による強制力を持つ規範。憲法、法律、条例などのこと。

【注釈】日本政府は、2022 年 4 月に策定された「AI 戦略 2022」を起点に、AI に関する政策と制度の整備を段階的に進めている。同戦略では、AI を用いた危機対応(災害・気候変動など)と社会実装推進が重視され、信頼性・安全性確保や人材育成も柱とされた。2023 年には、生成 AI の普及を受け、経済産業省と総務省が既存指針を統合する「AI 事業者ガイドライン」の策定を開始した。2024 年 4 月に第 1.0 版が公表され、AI 開発・提供・利用の各段階における基本原則が整理された。さらに、2025 年 6 月には「人工知能関連技術の研究開発及び活用の推進に関する法律」(AI 推進法)が公布され、政府の責務や AI 基本計画の策定、司令塔機関ともいえる「人工知能戦略本部」の設置が定められた。同年 9 月に戦略本部が発足し、初の AI 基本計画の策定が始動した。これら一連の動きにより、日本はソフトローとハードローを組み合わせた AI ガバナンス体制の構築に踏み出している。

【本章のまとめ】

- 2050 年では AI ロボットが目的を与えられて自律的に動くであろう。ただ 2050 年においても、AI ロボットが自分自身で目的そのものを設定するかは不明である。

また本項目では AI ロボットの自律性に関する論点を整理し、主に以下の論点を列挙して見解を述べた。

<倫理的課題>

- 事故発生時の責任の所在

<法的課題>

- AI ロボットの人格の問題
- AI ロボットの法的な規制
- AI ロボットが関連する知的財産権(論文ほか)の課題

<社会的課題>

- デュアルユースの問題
- 責任に関する社会的な受容性の問題

3. AI ロボットに関するデータガバナンスの ELSI 上の問題点について

【関連情報・キーワード等】

- 個人情報保護法、DFFT(Data Free Flow with Trust:信頼性のある自由なデータ流通)、デジタル行財政改革会議
- 欧州(EU): GDPR(General Data Protection Regulation), EU Data Act、EU Data Governance Act

【論点の整理】

AI・ロボットが人間とインタラクションする際に、収集・保存・利用されるデータの取扱いにおけるELSIをめぐる論点を整理する。目標3のプロジェクトにおいては、介護・看護・医療等の「ケア」領域へのAI・ロボットの適用を想定している研究(注釈)があり、その実証実験や社会実装においては、「ケア」の質の向上のために個人情報を含む様々なデータが収集・保存・利用されることがありうる。かつて個人情報については、濫用等の不適正利用からの情報主体の「保護」を念頭において議論がなされてきたところであるが、個人情報保護法において「個人情報の適正かつ効果的な活用が新たな産業の創出並びに活力ある経済社会及び豊かな国民生活の実現に資するものであることその他の個人情報の有用性に配慮」すること(同法1条)が明記されているように、現在においては個人情報の適正な利活用を可能とすることも重要な論点となっている。我が国においても、「プライバシーやセキュリティ、知的財産権に関する信頼を確保しながら、ビジネスや社会課題の解決に有益なデータが国境を意識することなく自由に行き来する、国際的に自由なデータ流通の促進を目指す」というDFFT(Data Free Flow with Trust)の推進や、デジタル行財政改革会議によるデータ利活用の促進などの取組が進められている。同様に、EUにおいても、従来の個人データ保護を主眼とするGDPR(General Data Protection Regulation)に加え、任意によるデータ共有の促進を意図してデータの処理手続や構造を規定するData Governance Actや、IoT(Internet of Things)の重要性の高まりの中で、デジタル・トランスフォーメーションの促進に寄与することを意図して、価値あるデータの公平な分配を可能とするEU Data Actなどが整備されているところである。したがって、目標3においても、このような社会・産業のニーズの変化(Social Issues)に即した法構造の変遷(Legal Issues)を、その基盤となる倫理的課題(Ethical Issues)と共に把握し、技術の開発・社会実装へと反映することが求められることになる。

【注釈】目標3のプロジェクトのうち、菅野プロジェクト、平田プロジェクト、下田プロジェクト、それぞれの研究開発項目の一部が該当する。

<倫理的課題>【菅野プロジェクト、平田プロジェクト、下田プロジェクト、原田プロジェクト、牛久プロジェクト】

データの取扱いに関する倫理的課題を考察する上で最も重要な論点は、データをどのような財として理解するかである。この問題について、一方に、そのデータを生み出した活動の主体(「データ主体」)に帰属し、その主体が自由に処分できるものであるという考え方(「データ=私有財」)が存在する。例えば、個人データと人格の表出として理解するような場合には、こうした考え方に基づいて倫理的課題を理解することが素直である。また、産業としてデータを活用するためには、一般に適切なデータの収集と処理とが必要であり、投資によってデータを価値あるものとした主体が存在する場合には、そのような主体にデータが帰属するべきだという考え方にも倫理的直観に訴えかけるものがある。しかし、「データ=私有財」という考え方は、様々な個人データが統合されることによって

初めて社会的に価値のあるデータの利用が可能となるという局面や、特定の産業主体がデータを独占することによって、社会的により価値の高いデータの利活用の機会が失われる局面などが存在しうること、しかもデータには通常の私有財とは異なり排他性がないため、利用競合という問題が生じにくいという性質があることなどを勘案すると、常に適切であるとは言い難いように思われる。その一方で、データに関しては、様々なデータが連結されることで初めて、社会的により価値ある利用が可能である局面が少なくないため、そもそも広く公共で共有して利用されるべきであるという考え方も存在する（「データ＝公共財」）。例えば、医療やケアの局面において、難病の患者に対する有効な治療法やケアのあり方を検討するためには、その難病に関する情報が広く専門的研究機関によって共有されて研究が進められることが望ましい。もっとも、「データ＝公共財」という考え方についても、差別的取扱いにつながるようなデータが自由に流通することによって、データ主体が思わぬ形で不利益な取り扱いを受ける可能性がある（「モラル・ハザード」）という問題がある。これに加えて、相当の投資によって生成したデータであっても常に公共財として他者と共有することが求められるとすると、その価値に誰も気づいていないデータについて先行投資することで、新たなビジネス機会を創出しようというインセンティブが減退する可能性がある（「フリー・ライダー」）という問題も指摘することができる。したがって、現時点においては、どちらかの考え方に立てば完全な解が得られるというわけではなく、むしろそれぞれの考え方が具体的な局面でどのような費用と便益とを生み出すのかに着目し、これを丁寧に衡量し、便益を最大化することが求められるといえよう。

＜法的課題＞【菅野プロジェクト、平田プロジェクト、下田プロジェクト、原田プロジェクト、牛久プロジェクト】

倫理的課題において指摘したように、データに関しては、「データ＝私有財」及び「データ＝公共財」という2つの考え方があるものの、どちらかが決定的に優位であるというわけではなく、結局のところ、それぞれの考え方が具体的な局面においてどのような費用と便益とを生み出すことになるのかに着目して議論を進めることが望ましい。一方で、従来の法的議論においては、いわゆるプライバシー権が憲法13条の幸福追求権との関係で、個人の人格権と結び付けられて議論されてきた経緯がある。そのため、「データ＝私有財」の考え方に傾斜し、データ主体に不利益が及ばないことを主たる目的として、法的制度のあり方が議論されてきた。しかし、このような法的議論の延長線上において、現行法制上重視されてきたデータ主体の「同意」のような方法が、複雑化するデジタル社会において、データ主体を潜在的な不利益から保護する上でどの程度有効なのかについては、再考の余地がありうる。一方で、とりわけ医療やケアの局面においては、個人に最適化した取り扱いを受けられるようにすることこそが、憲法13条の幸福追求権の保障にとって重要であるという議論もありうる。この方向からは、「データ公共財」のような考え方にに基づき、むしろデータの共有を進められるようにすることこそが、法制度としても求められるということになるだろう。もっとも、その際には、同時にデータ主体に差別的な取り扱いのような不利益が及ばないように、あるいは先行投資主体に不測の不利益が及ばないように、ガバナンス上あるいは産業育成上の適切な法的措置の検討が求められるところである。これらの問題については、現在日本も含む各国で議論が展開されつつある状況であることから、その議論の進展状況に目を配りつつ、技術の開発・実装に関する検討を進めることが期待されるといえる。

＜社会的課題＞【菅野プロジェクト、平田プロジェクト、下田プロジェクト、原田プロジェクト、牛久プロジェクト】

データに関する社会的課題としては、依然として個人データの不適切な利用・管理の問題や、産業機密に関するデータの持ち出しなどの不正競争に関する問題など、データ「保護」に関する問題群が存在している一方、少子高齢化や社会のデジタル化に伴い、価値あるデータを共有することで、データの利活用を通じた社会的便

益を最大化することも強く求められるようになっていく。これらの社会的課題をどのように解決するのかは、もとより日本のみならず国際社会においても大きな社会的課題となっている。医療・ケアの局面でロボットを利活用し、これらの社会的課題をどのように解決していくのか、その際、上述の倫理的・法的課題に対してどのように適切に対処していくのかは、研究開発の社会的意義を決定的に左右しかねない重要な問題である。そのため、ロボットの研究・開発の局面であったとしても、これらのデータの取扱いに関する社会的課題を解決することに貢献しうる設計や技術的措置のあり方についても、検討・議論することが望ましいと考えられる。

【注釈】「データ＝公共財」としての利活用は、医療技術などの発展や難病治療、科学技術の進展にとって重要である一方で、個人の尊厳と差別防止の観点からの法的制約が不可欠である。研究の公益性と個人の権利保障との適切なバランスをとる制度設計と運用が、今後ますます重要になる。

【注釈】日本の個人情報保護法は、個人データの「匿名化」に関して、匿名加工情報と仮名加工情報という2つの法的概念を規定している。匿名加工情報は2015年改正で導入され、特定の個人を識別できず、かつ元の情報に復元できないよう加工された情報を指す。加工後は、再識別の禁止、公表義務、安全管理措置などが課され、適正な基準を満たせば個人情報とは見なされず、本人の同意なく第三者提供が可能となる。一方、2020年改正により創設された仮名加工情報は、他の情報と照合しない限り個人を識別できない状態に加工された情報であり、主に事業者内部での利用を想定している。一定の条件下で利用目的外利用が認められるが、復元可能性を前提とするため、引き続き個人情報に該当し、第三者提供は禁止される。

両制度は利活用とプライバシー保護の両立を目指すものの、実務上の課題も指摘されている。匿名加工情報は制度自体が手続と基準を複雑化させ、かえって規制強化となっているとの批判がある。また加工基準の不明確さや、作成目的の主観的要件が制度運用を困難にしている。仮名加工情報も、想定された用途との乖離や第三者提供の制約などから、柔軟な利活用に限界があるとされる。こうした課題を受け、制度の見直しや新たな枠組みの必要性が政府内でも検討されている。

【本章のまとめ】

本章では、AI・ロボットが人間とインタラクションする際に、収集・保存・利用されるデータの取扱いにおけるELSIをめぐる論点を整理した。

< 倫理的課題 >

- データの取り扱いに関する倫理的課題

< 法的課題 >

- データの取り扱いに関する法的課題

< 社会的課題 >

- データの取り扱いに関する社会的課題

4. 月面での活動における ELSI 上の問題点は何か

【関連情報・キーワード等】

- 宇宙条約、宇宙物体登録条約、月協定、アルテミス合意、宇宙空間平和利用委員会 (COPUOS)
- 宇宙研究委員会 (COSPAR: Committee on Space Research) の惑星保護指針

【論点の整理】

<背景>

宇宙空間(天体を含む)の領有自体は、すべての宇宙活動国が加盟する宇宙条約(1967 年)により明確に禁止されているが、長期の占有や排他的な利用自体は禁止されていない。また、宇宙資源の開発・利用・所有が自由なのか規制が必要なのか、そして規制の強度はどのようなものであるべきかについては各国の見解は一致せず、現在不明瞭である。この状態を解消すべく 2017 年から国連の宇宙空間平和利用委員会(COPUOS)の法律小委員会で開始され、2021 年以降は作業部会での集中的な議論が行われている。2027 年以降、第1段階のガイドラインが国連総会で採択される予定であるため、完全な自由競争とはならず、何らかの国際枠組・規制が勧告される可能性が高い。

アルテミス計画などによれば、2030 年代に月に拠点構築が始まる計画があり、2050 年には月拠点に人間が活動している可能性がある。ちなみに中国やロシアが主導する「国際月面研究基地(ILRS: International Lunar Research Station)」は、2035 年を目途として建設を目指している。

2050 年の月拠点内の人間の活動において AI ロボットが共存している場合には他の 2 章や 3 章と同様の論点が生じる見込みである。

2035 年頃には月拠点を構築するための探査活動やその後の建設活動が行われるフェーズであり、現在の目標 3 の国井プロジェクトと吉田プロジェクトはまずそこが ELSI の対象となる。

吉田プロジェクトでは究極的にはロボット自身が判断して活動することをゴールにしている。活動する場面は人が月に行く以前の段階もある。拠点ができた以降もメンテナンスでロボットが関わるので、物理的に同じ場所ですとロボットが共存することはある。その場合の ELSI は地上での共存の考えを参考にする。

国井プロジェクトでは拠点構築も目的になっているが、人を住まわせることは目的外と考えている。国井プロジェクトのシナリオでは人と一緒にロボットを使わないが、もし別の組織がこのロボットを使いたいといった場合に人と一緒に使う可能性もゼロとはいえないので、使い方次第だと思われる。

<倫理的課題>

月の環境保護の倫理課題【吉田プロジェクト、国井プロジェクト】

AI ロボットが月面で活動を行う場合に、月の環境に何らかの影響を与えるとすれば、環境保護の観点で倫理的な課題が生じるリスクがある。目標 3 のプロジェクトの AI ロボットについて、現時点での認識を述べる。

国井プロジェクトでは、探査ロボットにそれほど(環境を変える)力はない。仮に溶岩チューブに人が住めることになり、企業が大規模に開発するとなると景観も大きく変わるだろう。ただアルテミス計画には月面に拠点を多く作る図があるので、おそらくその段階で議論がされているものと思われる。

吉田プロジェクトでは、環境配慮は必要だが、アルテミス合意の大きな流れに従って計画が進展していくと理解している。構造物を作るので月面に変化を与えるため、もし景観を破壊しているといった議論が出てきたら国際的な議論の中で対応していくものと認識している。

宇宙倫理学における景観に関するリスク【吉田プロジェクト、國井プロジェクト】

月面に関する目標3の現在のプロジェクト(吉田プロジェクト、國井プロジェクト)の研究開発範囲においては、倫理的な問題は現時点ではあまり見当たらないと思われる。ただ、大学や研究機関には宇宙倫理の研究者も複数おり、その内容は参考情報として認識しておいた方がよい。例えば、京都大学宇宙総合学研究ユニットの先生方が作成した「宇宙倫理学」の教科書や、稲葉辰一郎先生(明治学院大学)の「宇宙倫理学入門」の本などがある。関係する話としては、火星に行けば生態系で気にしないといけなくともある。「景観」というものが環境倫理では対象になっている。月の景観(ランドスケープ)を変えるようなことについては配慮が必要である。宇宙空間に広告を出すというのも景観を壊す。宇宙や月は文化的に重要なものなので倫理の対象になる。必ずしも「景観」維持という環境倫理を追求したものではないが、宇宙空間での広告をどこまで許容するかについては国連宇宙空間平和利用委員会(COPUOS)の科学技術小委員会(科技小委)で議論されたことがあり、また、米国法典 51 編「国家・商業宇宙プログラム」では、「ひどく目立つ広告」(obtrusive space advertisement)目的に利用される材料を含む物体を打ち上げることができない旨を商業宇宙打上げ法で規定している(51 USC § 50911)。

また宇宙空間への人間の進出については、人間のサイボーグ化を危惧する考えもある。もともと「サイボーグ」の言葉を考えた人たちは、宇宙に適用するために人間を改造したほうが効率的ではと考えた。なお目標3の AI ロボットの中には人に装着するタイプのものがあるが、「サイボーグ化」といえるものではないと考える。目標3の現在のプロジェクトの AI ロボットが、月面の景観を変えるほどの活動をするとは考えにくいために景観リスクは考えにくい。将来の開発においては景観リスクが発生しないように留意することは必要ある。なお本章の＜社会的課題＞の項では「月面に放置される AI ロボット」として、社会的な課題の可能性について述べている。

宇宙活動における宗教観の課題【吉田プロジェクト、國井プロジェクト】

宇宙空間には上下の感覚がないなど地球上の環境とは大きく違うため、人間関係の上下関係などの価値観をはじめとして、宗教観や倫理観も変わるのではないかという研究もある。月面上においては、月面上と上空など多少の上下感覚はあると推測するが、地球上との違いがどのように影響するかは不明である。現時点の目標3のプロジェクトの AI ロボットの探査活動や拠点構築の初期段階では、人とのインタラクションはほとんど無いと想定されるが、物理的に同じ場所で人とロボットが共存することはありえる。月面上で、人間の数が少ない環境(初期は数人から十数人程度かもしれない)の中で、複数(数個から十数個)の AI ロボットの活動を眺めながら、上空の暗闇の宇宙空間を眺めながら、果たして人間はどのような感覚になるのだろうか。寂しさなど、諸々の感覚が生じて、宗教的な救いを求める場合もあるかもしれない。その場合、AI ロボットの外見の形状や動作や機能が何らかの影響を与える可能性もあるかもしれない。

＜法的課題＞

月の環境保護の法的課題【吉田プロジェクト、國井プロジェクト】

現時点で法的には、ほぼ適切な対応がとられているという認識である。

探査の場合には、惑星検疫等の宇宙研究委員会(COSPAR:Committee on Space Research)の規則を守るということで問題は無いと思われる。

国家管轄権の適用の問題(どこの国の法律が適用されるかの問題)【吉田プロジェクト、國井プロジェクト】

目標3のAIロボットを月に打ち上げる際には、宇宙物体として、AIロボットそのものと、打ち上げ国や所有国について登録が必要になる。ただし、部品で登録するか全体で登録するかは決まっていないというのが実態である。宇宙物体は、これまでは主として衛星が登録対象だったので99.9%はあまり考えずに済んだ。例えば、日本の「かぐや(月周回衛星)」から「おきな」と「おうな」の子衛星が出る場合、登録としては「おきな」と「おうな」も登録はしたが、打ち上げは1つの打ち上げと登録した。登録簿の書き方が何年の何番目という登録で、1つの宇宙物体が分離したという登録になっている。ロシアは軍事衛星からさらに衛星を放出した場合に、新しい打ち上げ行為として別の登録をしているケースもある。また、一つの衛星から探索用の機械を出すときに登録しない国もある。衛星の分離と捉える場合、新しい打ち上げ行為と捉える場合などがある。月面での探査が始まると何らかの整理が必要になると認識している。

國井プロジェクトの場合は、ロボットをランダーに搭載する際に、ランダーの一括りで登録するのか、ランダーからロボットを分離して登録するのかといった登録方法については必要な手続きに従うが、具体的な登録の際に何らかの課題が出てくる可能性はある。

吉田プロジェクトの場合は、ロボットのモジュールを分割して、多くのモジュールを運び込む。また過去に運んだモジュールと新しく運んだモジュールが結合して新たなロボットを作る場合もある。これらの物体の登録についてはガイドラインに従うが、具体的な登録の際には登録の仕方に課題が出てくる可能性がある。ただし、登録に基づく国家管轄権の行使が本当の意味で問題となるのは、多国籍の民間企業がさまざまなロボットや機器を用い、その売買やリースなどが開始して以後のことであろう。国家機関の地位の有無を問わず、宇宙活動国の宇宙機関や、宇宙機関からの委託を受けた同一国籍の私企業が探査活動を行う時期には、アルテミス合意の参加国同士、ILRS 参加国同士の国家間協定に基づいて活動がなされるため、国家管轄権問題は前面に出てくることはほとんどないだろう。例外は、アルテミス合意参加国とILRS 参加国との間で管轄権行使問題が生じるときであろうが、この場合も登録に基づく管轄権というよりは、月の場所の使用権、資源の採掘権の争いのような形を取るとと思われる。

リスク管理・責任分界点・公共危険など【吉田プロジェクト、國井プロジェクト】

地上のロボットのリスク管理のやり方では遠隔操作で区分することが多いが、目標3の現在の2つのプロジェクトは最初からすべて遠隔操作が想定されているので、異なる区分が必要と考える。人とのインタラクションが特にないが、他のシステムとの協調など、システムのリスクの形で管理の仕方が変わる局面を考えるのが良いのではない。システムが複雑になってくると、リスク評価の仕方が変わってくる区分があるのではないだろうか。

月面でロボットのある種のプラットフォームのような物理的なプラットフォームを作り、それと協調しながらロボットがいろいろな活動することが将来は起きるだろう。そのプラットフォームは共用施設として基本的には国際的な協力で作る話になるだろう。それを破壊することは公共危険のような話になる可能性があり、それについては法的には

違う取り扱いをする。地上での（電信柱に車が突っ込み信号に影響するような）事故は比較的復旧が早い、月面での事故は質的に違う問題が出てくる可能性があるだろう。そのため月面でのリスクは、地上での想定リスクとは区分けが違い、重み付けなども考えた方が良いのかもしれない。

宇宙では無人システムと有人システムは、はっきりと線引きされている。無人システムは間違って壊しても人が死ぬわけではないが、国際宇宙ステーションのような有人システムになると、一つ一つのものが、そこに滞在する宇宙飛行士の生死に関わるので細かく決められており、しっかりとした体系の中で動いている。一方で無人システムの人工衛星に関しては間違って壊しても、その衛星を所有してる人が困るだけの話となる。目標3の現在のプロジェクトの AI ロボットは無人システムの部類と想定されるので、その AI ロボットの故障・暴走したとしても直接的に人の生死には影響しないであろう。ただし、もし月面の共用施設に人間が滞在している状況において、無人システムの部類の目標3の AI ロボットが故障・暴走したことで、月面の共用施設のインフラに影響を与えて壊してしまい、それが共用施設の人間の生命維持装置に万一影響を与える場合には人命リスクが起きる可能性もゼロではないだろう。なお、共用施設の一例として通信設備（月面における通信、月面と地球との通信など）が想定される。初期の段階では通信設備の数は少ないため、それが故障すると影響が大きい、月面開発が進んだ段階では、月面上の通信設備の数も増えているものと予想され、そのような時代になれば、地上と同様に一カ所が壊れても他の代替え設備で利用できることも期待される。

宇宙の責任体制は、全てが公共危険のような形で扱われているとも言える。細かいルールはないが、その一つの証拠になると思うのは、宇宙条約の中の6条である。普通、国家は国家機関の行なった国際違法行為に対して外国に対して責任を負うが、全ての私人の活動が国家機関の活動と同じように見なされて、国が責任を負うという形にしている。これは、宇宙空間の特殊性というところから来ている。ただ、宇宙条約第6条は、原則であり、個々の例において直接適用できるほどの具体性をもった規則とは言えない部分がある。実際に国家機関ではないという意味で私人であるテロリストが外国領域から外国の衛星を破壊した時にどういう形で国が当該外国に対して責任を負うのかなどの実例がないので、結局は普通の国際違法行為責任とどう違うのかというところはまだまだ未成熟といえる。だが、原則という面では、公共危険のような考え方で作られていると言っていいと思う。

損害の責任【吉田プロジェクト、國井プロジェクト】

今のところ宇宙物体が地上に物理的損害を及ぼした場合には、無過失で上限なしの賠償責任になっている。天体を含む宇宙空間では過失責任となる。しかし天体のようなところでの活動が盛んになった時に、今の過失責任で行くのが合理的なのかどうかという論点も生じるであろう。

ロボットや AI に関していえば、地上でも過失責任でいけるのかという議論自体がかなり活発になされ始めていると理解しているので、ますますややこしいポイントになるだろう。特に複雑なシステムが組み合わされて起きるリスクでの過失責任は扱いが困難になり、今回のプロジェクトの群知能の話が関連してくるだろう。

企業の国籍国と被害を受けた企業の国籍国の両方が宇宙損害責任条約に入っている場合には、国対国の交渉で損害責任を過失に基づいて問う。具体的には、1つの宇宙物体が別の宇宙物体に対して物理的な損害を加えた場合、それぞれの物体の「打上げ国」の過失の有無に応じて損害賠償責任の有無と損害額の決定を行うことになる。ここで問題となるのは、「宇宙物体」が、宇宙空間に導入される地上の人工物の総称とされていることである。仮にロボットが地上から打ち上げた宇宙物体と天体の資源との双方を用いて現地で製造された場合には、これは宇宙物体の定義に合致せず、したがって、宇宙物体の打上げ国同士の交渉により損害賠償を決定するという仕組みを用いることができない可能性が生じる。そのような法的不明瞭を考えると、今後作っていくルールは多

いといえる。

損害責任は打上げ国が連帯責任として負い、登録国は打上げ国の部分集合なので宇宙物体を登録した国は必ず打上げ国責任を負う。物体の所有国はそれ自体では必ずしも打上げ国とはいえない。例えば打上げ国から購入した国やその私人の国籍国である場合がある。月の活動の初期には、国家が関与するのでその物体の打上げ国は必ず登録をするため、管轄権を行使し、打上げ国のうちの一国である登録国が明確になると考えられる。問題が生じるのは私企業の活動が活発になってからのことであろう。

月面活動における知的財産権の課題【吉田プロジェクト、國井プロジェクト】

現時点の目標3のプロジェクトの AI ロボットは、地球上(大半は日本国)で発明され、国内および外国(複数国)に出願・登録される。一般に、その特許権の効力は取得した国の領域内に限られるという「属地主義」がある。特許権の侵害についても属地主義が関係する。過去の事例として、国際宇宙ステーションは複数の国が共同利用しているが、宇宙空間において発明がなされた場合の知的財産権の取り決めについては、「国際宇宙ステーション協定」というルールが作られた⁴⁾。今後、月面上において発明がなされた場合の知的財産権や侵害については、どこの国の法律を使うかの課題があり、国際法上の合意が必要になるであろう。つまり、目標3の AI ロボットが月面上で活動し、中期的な月面上の活動において AI ロボットに関する発明がなされた場合、どこの国の法律を適用するかが課題となる。なお、フランス、ドイツ、米国では知的財産法の適用範囲を宇宙物体上にまで拡大しているが、日本では適用拡大されていない⁵⁾。すなわち、目標3の AI ロボットが宇宙物体として登録されて月面上で活動する状況においては、属地主義により、日本及び外国で取得した特許権の効力は発生しないため、他国がその特許を月面上で侵害しても法的対策は取れないと思われる。今後の宇宙活動の進展を鑑みれば、日本の特許法の宇宙物体上への適用範囲拡大が必要と思われる。

一方、AI ロボットが月面上で撮影した画像については、機械的に取得した画像の「一次データ」として創作性がなく著作権に該当しないと思われるが、その画像に対して何らかの処理や解析を行った情報に対しては著作権上の保護が与えられる可能性は考えられる⁶⁾。

一般的な議論ではなくて、目標3の吉田プロジェクトおよび國井プロジェクトでまずもって重要なことは、両プロジェクトが開発を進めている、さまざまな技術について国際特許を取っておくことである。

<社会的課題>

月面に放置される AI ロボットの扱い【吉田プロジェクト、國井プロジェクト】

月面の探査や拠点構築において活動した AI ロボットはその活動を終わると月面もしくは溶岩チューブ内に放置される状態になるか、あるいは何らかの手段で回収されることになる。放置される場合には、その AI ロボットの物体を「ゴミ」と考えるか、「科学遺産・文化遺産」と考えて保護するという考えもある。数量が増えればスペースデブリ問題になるだろう。アルテミス合意でも科学遺産・文化遺産の規定はこれから作り上げていくと述べられている。物理的なゴミ問題は、COSPAR というより、IADC (宇宙機関間デブリ調整委員会: The Inter-agency Space Debris Coordination Committee) や各国の基準によるのかもしれない。この問題は今のところルールはない。非常に難しい問題で、法がしっかりあるわけではない分野である。大事なものは、探査の価値と他の価値との間でぶつかる場合は、比較考慮して何をどう重視していくのかという根本的な軸を考えることである。具体的な事例に沿って今の活動を阻害しない、しかし他者に与える損害も少なくする。具体的なことと、もう少し大きなことの両方を考えていく

必要がある。

月面活動の衝突のリスク【吉田プロジェクト、國井プロジェクト】

目標3の國井プロジェクトのAIロボットの探査場所は関係者の協議次第と思うが、他国と重複する可能性はある。また吉田プロジェクトの月面の場所については、有人拠点を目指すので極地域を狙う。月の水資源で国際的な注目を集めている場所なので注意を払う。基本的にはアルテミス合意の中で役に立つロボットを作っていくが、非アルテミスのグループとの調整が発生すると予想される。

このように、アルテミス合意側の国と、中国・ロシアが主導する国際月研究基地(ILRS)側の国の衝突があるかもしれないと懸念している。

月面の領土獲得自体は禁止されているが、今のルールでは占拠は禁止していない。占拠することによる主権の主張が禁止されている。目標3の月面のAIロボットの活動を行うために必要な「広さ(面積)・三次元的な高さ(太陽光パネルなど関連の設備を含む)」や「活動する期間」などについて情報を整理することで、場所の調整リスクの度合いが具体的に見えてくると思われる。

宇宙条4条では、天体上では活動は「もっぱら平和的目的」(exclusively for peaceful purposes)で行うことが規定される。平和的目的の一般的な解釈は、非侵略的な＝自衛権の範囲内の軍事利用を含むというものなので、これだけであれば、天体上では一定の軍事利用はできそうであるが、あらゆる基地の設置・軍事実験・軍事演習が明示的に禁止されているので、実質的には、非軍事利用が課されていると解されている。宇宙空間では、禁止されているのは大量破壊兵器の設置のみであり、天体上の厳しい軍備管理とは異なる。宇宙空間でも天体上でも、武力攻撃を受けた国が、対抗して取る自衛権の行使としての武力行使は禁止されていない。今後、欧州・米国・日本などのアルテミス合意国と、中国・ロシアなどのILRSのグループ側とが、それぞれ独自に月面活動におけるルールを策定するような事態になった場合は、月面上の活動が衝突するリスクがある。

2025年6月時点の現状について説明する。中国・ロシア側の活動に関しての特筆事項として、中国が主導し、ILRSを国際機関とし、安徽省に本部を置いた。アルテミス合意側も月資源の探査・利用を進めようとしている。アルテミス合意内容の約80%は宇宙関係条約の焼き直しであるが、資源探査・利用や宇宙遺産の保護など新たな分野にも触れている。アルテミス合意の規定する月面活動に従事する者の安全区域の設置・利用規則はILRS側を法的に拘束するものではなく、ILRS側は、別のルールを作って行動する可能性がある。共通のルールがないことから、将来的に両者が同一区域での活動の正当性を主張して衝突する可能性がある。2017年に国連宇宙空間平和利用委員会(COPUOS)法律小委員会で宇宙資源の探査・開発・利用に関する法規範を議論するアジェンダが立ち上がり、2021年から宇宙資源ワーキンググループの名称で作業部会が設置された。2027年の国連総会決議採択を目指して、2025年5月に草案のドラフトが出たが、まだまだ収斂する方向に向かってはおらず、2029年頃の国連決議になる模様であり、その時点でも不明瞭性が残る見込みである。

宇宙条約2条では領有取得は禁止されているが、占有自体は禁止されていない。資源の利用は禁止されておらず、利用の中に所有まで含めるかどうか意見の分かれる部分である。日本は、許可を得る条件の中で国際法と合致することを前提に法律を制定しており、仮に国際法の利用の中に所有権の主張が入らなくなると進め方も変わる。アメリカは国際規則に従って所有権を自国民に認めるというものであり、自国民への所有権認定の考え方が異なる。ルクセンブルク、アメリカ、日本のような国は、宇宙の利用の中に資源の所有が入るということを前提に探査を進めている。探査段階を進めて開発段階に持っていくときに安全区域をどう作っていくか、その場合の協議の具体案をアルテミス側は出している。それは今後さらに精緻化される見込みであり、一方でロシア・中国側も

明確化してくると思われるので国連内外の場で議論に参加していくことが大事になるだろう。南極の場合でも日本は早くから探査を行っていて協議会に入っているため、第一グループに入るための探査を進めていくことが大事だと考える。

【用語説明】

アルテミス合意:宇宙空間の探査・利用のガバナンスのための共通ビジョン。

スペースデブリ:軌道上にある不要な人工物体のこと。運用を終えた人工衛星、故障した人工衛星、爆発や衝突により発生した破片などのこと。

【本章のまとめ】

- 月打ち上げに関する宇宙物体の登録についてはプロジェクトの内容に応じて登録の仕方を検討する必要がある。
- 月面の AI ロボットの活動は現時点では遠隔操作と自律動作によるものと想定しており、人とのインタラクションはないため直接的な人命リスクは考えにくい。しかし、月面に国際的な協力により共用施設のインフラや AI ロボットとの連携のためのプラットフォームが作られた場合には、AI ロボットの故障・暴走により、その共用施設に影響を与える可能性があり、法的な「公共危険」の考え方が適用される可能性がある。
- 月面活動のルールは定まっていない部分も多い。最終的には政治的な判断になるのかもしれない。正式な国際法としての条約は国連では予見可能な将来、できないかもしれない。最近重視されているのはオペレータの側（行動している共同体）の最も良いルールが何かである。それを政府等に提示してルール形成のための行動をとってもらおう。倫理面でも使いやすさの面でも、どういうルールがあると便利なのか、どういうものは困るのか、対外的に提案できるものが出てくれば役立つと思う。

本章では以下の論点について、それぞれ見解を述べた。

<倫理的課題>

- 月の環境保護の倫理課題
- 宇宙倫理学における景観に関するリスク
- 宇宙活動における宗教観の課題

<法的課題>

- 月の環境保護の法的課題
- 国家管轄権の適用の問題（どこの国の法律が適用されるかの問題）
- リスク管理・責任分界点・公共危険など
- 損害の責任
- 月面活動における知的財産権の課題

<社会的課題>

- 月面に放置される AI ロボットの扱い
- 月面活動の衝突のリスク

5. 社会的な受容性の観点

【関連情報・キーワード等】

- 日本の人口減少、高齢社会
- 格差、不平等
- 不気味の谷

【論点の整理】

2030 年における AI とロボットの状況と、2050 年における AI とロボットの状況とを分けて AI とロボットの社会的課題を考える。

【注釈】将来の状態(2030 年における AI とロボットの状況、2050 年における AI とロボットの状況)を想定(想像)することは、かなり大変であるが、考えておくことは重要である。将来の状態の予想に、現状によるバイアスが働くことは否定できないが、できるだけバイアスが働かないようにして考える。

<背景>

日本では人口減少が進んでおり、高齢化率(65 歳以上人口割合)も増加していく。内閣府の令和 7 年版高齢社会白書⁷⁾によると、総人口は、令和 6 年 10 月 1 日現在、1 億 2,380 万人(高齢化率は 29.3%)である。将来の予測は、2030 年で総人口 1 億 2,012 万人(高齢化率は 30.8%)、2050 年で総人口 1 億 469 万人(高齢化率は 37.1%)である。一方、国際連合広報センターの人口と開発の掲載情報⁸⁾において、国連の推計によると 2015 年の世界の人口(年央推計)は 73 億人で、2030 年までに世界の人口は 85 億人に達し、2050 年には 97 億人に増加するものと予測されている。こうした増加のほとんどがアフリカとアジアで発生すると思われると記載されている。

日本では現在でも各業界で人手不足が課題となっており、2050 年に向けてさらに人手不足が顕著になることが予想される。令和 7 年の現在においても人手不足の解消のために各業界では様々な取り組みを行っている。例えば、日本に入ってくる外国人労働者の活用や 65 歳以上の高齢者の活用といった人による対策がある。またデジタルの活用(デジタルトランスフォーメーション:DX)などによって業務効率化などを行う対策である。後者には生成 AI による業務効率化や一部の業界へのロボット導入などの試行や実践も行われている。

目標3では人手不足解消の手段として、将来には AI ロボット導入を目指しているが、導入においては社会的な受容性が課題となる。社会的な受容性は世界各国の宗教や文化等にも関連することが容易に推定されるため、まずは我が国の日本における受容性を議論する。もちろん日本においても地域差や個人ごとの文化や価値観の違いがあるが、ひとまず日本全体で共通的に考える。

ムーンショット目標3は 2050 年の社会に向けた研究開発であり、2050 年において AI とロボットが人と共生している状況を目指している。つまり 2050 年において目標3が思い描く AI ロボットが社会に広まっている想定である。その状態において存在すると思われる社会的な課題を考える。一方、2030 年においては目標3の AI ロボットが社会に広まっている状態ではなく、おそらく社会実装のための実験や一部の限定的な社会導入に留まっていると思われる。もちろん最近の世界の研究開発の加速状況を見ると、2030 年においても AI とロボットが社会の中でかなり導入されている可能性もあるが、ここでは実験レベルや一部限定導入の想定で考えることとする。

＜2030 年における社会的課題＞

2030 年の日本社会

2030 年の日本は総人口約 1 億 2,012 万人、生産年齢人口（15～64 歳人口）は約 7,076 万人、高齢化率 30.8% となる見通しである（国立社会保障・人口問題研究所『人口統計資料集（2025）』⁹⁾）。出生数は年間 70 万人台で推移し、合計特殊出生率は 1.30 前後で推移する（国立社会保障・人口問題研究所『日本の将来推計人口（令和 5 年推計）¹⁰⁾）。世帯総数は 5,773 万でピークに達するが、平均世帯人員は減少傾向が続き、単独世帯が 41.6% を占める（国立社会保障・人口問題研究所『日本の世帯数の将来推計（全国推計）』¹¹⁾）。生産年齢人口は 2020 年から約 430 万人減少し、女性・高齢者の労働参加率上昇を加味しても、労働力の供給不足が約 341 万人に達するという推計もある（リクルートワークス『未来予測 2040』¹²⁾）。

これらの数値は、2030 年の日本が本格的な人口減少社会の課題に直面する転換期であることを示唆している。深刻化する高齢化や労働力の減少といった構造的な課題に対し、解決策の一つとして AI ロボットの活用が期待されている。しかし、その技術を社会に広く浸透させるためには、次のような社会的な受容に関する課題を乗り越える必要がある。

なお、目標 3 の中心的な論点ではないものの、生成 AI の普及にともなう電力需要の拡大も重大な社会課題である。生成 AI の普及は、その計算基盤であるデータセンターの電力消費を急増させ、日本のエネルギー需給と環境政策に新たな課題を突きつけている。三菱総合研究所の試算では、生成 AI の活用が最大限に進んだ場合、2040 年の ICT セクターの電力需要は 2020 年比で最大 27 倍に達する可能性が指摘されている（出典：三菱総合研究所『生成 AI の普及が与える日本の電力需要への影響』2024 年¹³⁾）。

このほか、階層の再生産・固定化や正規雇用・非正規雇用の賃金・待遇格差、ジェンダー格差、地域間格差など、社会経済的な格差・不平等も重大な社会課題であることはいうまでもない。21 世紀に入って日本社会では格差（所得格差など）は拡大してきている。格差縮小をめざした適切な対策が講じられないと、2030 年には、格差が拡大することが予想される。格差の拡大がもたらす社会の不安定化と地球環境問題は、これからの日本社会がかかえる重要な社会課題だと言える。ただし、以下では目標 3 と関連する論点に限り、AI ロボットの普及に伴う社会課題を論じる。

雇用・労働への影響【プロジェクト共通】

目標 3 では AI ロボットを導入することにより、生産性の向上や人手不足の解消につながることを目指して研究開発を行っている。

経営者（例えば介護施設）は人手不足解消のためにロボットの導入を考えるが、働いている人はロボットの情報が少なく抵抗感（仕事が奪われる、機械の苦手意識など）がある。この抵抗感を払拭するためには、ロボットが自動的に教えてくれるなど、ロボットも含めたシステムとして解決するしかないのではと思われる。社会実装の実験においても、この抵抗感を払拭するような工夫が必要であるが、実験段階ではシステム全体が整っていない可能性があるため、実際の機械を使うだけでなく、今後の全体像に関する十分な説明と理解を得ることが重要となる。

事故発生に関する不安【プロジェクト共通】

新しい技術の社会導入には人の抵抗感が起きやすいので、まずは誰もが賛成できるような実績を積み重ねながら社会的な議論を並行して進めるやり方になると思われる。社会導入の実験の初期段階で社会的弱者（幼児、

高齢者など)の事故が発生してしまうと、そこで社会的な拒否反応が出て導入が止まってしまうので、初期段階での事故が発生しないよう慎重な準備と実施が求められる。AI ロボットの事故発生率よりも人間のミス発生率が高くても人間の場合は許容されるが、AI ロボットは許容されない。

目標3の AI ロボットの適用場面として、人の健康や介護に関する場面も想定されており、技術的な事故回避の方策はもちろんのこと、利用者への十分な説明と万一の事故発生時の合意が必要となる。

AI ロボット活用の格差・不平等の発生リスク【プロジェクト共通】

2030 年はまだ十分に AI ロボットは普及していないと思われるが、反対に普及していないからこそ、AI ロボットの活用の先行者利益が生じている可能性もある。特に AI 単体としては、AI と協働して論文を作るというように、その AI を利用出来る人と利用できない人との差が生まれる。なお企業・大学・個人等においては、AI ロボットの導入に必要な資金を捻出できるかどうかといった経済的な格差も存在する。

また、AI が学習するデータに偏りがある場合(例えば、白人／男性重視の考え方、人間中心の考え方など)に学習の結果にバイアスがかかり、偏見が生まれるリスクも想定される。

さらに AI ロボットが普及することにより、仕事で AI ロボットを扱う労働者は新たに必要なスキルを向上させる必要も生じる。仕事の現場の人に AI ロボットを使うスキルがないと、企業内の AI ロボット導入推進部門や AI ロボットの開発企業が現場に出向いて指導などの対応をする必要が出てしまうため、逆に非効率となる可能性もあるので、労働者および従業員におけるスキルの学習が経営者にとっては重要な課題となる。

大学の研究においても AI が活用されることで、膨大な量の論文を書ける人とそうでない人の格差が生まれるであろう。また、その膨大な量の論文の査読をしないといけいない研究者も、査読のための AI 利用の可否が業務の効率の格差を生むであろう。

研究者の研究方法は、AI サイエнтиストと対話しながら論文を書く手法に変わるだろう。論文を書く AI サイエнтиスト、査読する AI サイエнтиスト、それぞれの立ち位置が違ういろいろな AI サイエнтиストがいる。新たな論文を書こうとするときに、一緒に協働していくものや、レビューするものなど、モデルや何を価値基準として振る舞うのかが変わっているものが複数いるイメージを持っている。

一方、AI ロボットが普及し始めた段階においては、その AI ロボットの活動環境の中で人が働くときに、周囲に人間が少ない職場環境であれば、孤独感が生まれる可能性もある。

<2050 年における社会的課題>

2050 年の日本社会

2050 年の日本は、「人口減少社会」がより深刻化していると予想されている。総人口は1億 469 万人まで減少し、生産年齢人口(15～64 歳人口)も約 5,540 万人となる一方で、高齢化率は 37.1%に達する(国立社会保障・人口問題研究所『人口統計資料集(2025)』⁹⁾)。出生数は年間 60 万人台前半まで落ち込むが、2030 年代と同じく合計特殊出生率は 1.30 前後で推移する(国立社会保障・人口問題研究所『日本の将来推計人口(令和 5 年推計)』¹⁰⁾)。世帯総数は 2020 年よりも 310 万世帯減少し 5,261 万世帯となり、平均世帯人員は 1.92 人、単独世帯が占める割合は 44.3%となる見通しである(国立社会保障・人口問題研究所『日本の世帯数の将来推計(全国推計)』¹¹⁾)。生産年齢人口は約 5,540 万人となり、2020 年から約 1,969 万人も減少する(国立社会保障・人口問題研究所『人口統計資料集(2025)』⁹⁾)。2040 年時点で約 1,100 万人の労働供給が不足するとの予測もあり、深刻な状

態が続く見込みである(リクルートワークス研究所『未来予測 2040』¹²⁾)。

このように、2050 年の日本は、人口構造の激変がもたらす「社会機能の維持」という根源的な課題に直面しているものと思われる。2020 年から約 2,000 万人も減少した生産年齢人口では、深刻な労働力不足を解消できず、経済成長どころか社会インフラの維持すら困難になりかねない。そのため AI ロボットの活用は選択肢の一つではなく、必須の選択となっている可能性がある。AI ロボットの社会的受容を目指すならば、現時点から以下のような社会的課題を検討するのは極めて重要であるといえる。

雇用・労働への影響【プロジェクト共通】

2030 年と同様に 2050 年においても、AI ロボットの導入のメリットは、生産性向上や人手不足への貢献につながるものと想定している。

2050 年の日本の少子高齢社会の状況においては、AI ロボットの導入に伴う人間の雇用への影響を論じている場合ではないかもしれない。2050 年の日本の経済状況によっては、外国人労働者が日本で働くことに魅力を感じず、十分な外国人労働者の人数が日本に来てくれないかもしれない。外国人労働者が魅力を感じない日本であれば、20 歳代や 30 歳代の若手労働者も日本で働かずに海外で働く人数が増えているかもしれない。また 65 歳以上の人が労働力として多く働くことを余儀なくされている可能性もある。その 65 歳以上の人が働くのを身体的あるいは精神的に支援するために AI ロボットが使われている可能性もありえる。この場合には、AI ロボットが人の雇用を奪うという影響ではなく、65 歳以上の人が働くことを助けるために AI ロボットがなくてはならない必要なものになっている状況もありえるかもしれない。

<参考情報>

内閣府の「令和 6 年度 年次経済財政報告」¹⁴⁾によると、「2023 年 10 月末時点で、我が国の外国人労働者数は約 205 万人と過去最高を更新し、全雇用者の約 3.4%を占めるまでにプレゼンスが高まっている。現在、製造業・非製造業ともに企業の人手不足感が歴史的に高い状態にあり、少子高齢化に伴う労働力人口の下押し圧力が続く中にあることは、今後も労働市場における外国人労働者の重要性は高まりこそすれ、低下することはないと見込まれる。」とある。

事故発生に関する不安【プロジェクト共通】

2050 年における社会において一定程度の AI ロボットが導入されていると想像した場合、人の代わりに AI ロボットが何かの作業を長期間に渡り 100%完全に遂行することは難しく、ごく低い割合でも事故が発生する可能性がある。AI ロボットが事故を起こす割合(事故発生率)が、人が行う場合の発生率よりも低い数値であったとしても社会的に許容されない可能性がある。2050 年においては、AI ロボットが導入される業界や生活場面が増えていくことが期待される。AI ロボットを導入することで、全体で事故の件数を減少させるというのは大前提となるが、それをクリアしたらロボットを入れた方が良いという単純な話ではないだろう。得てして、テクノロジーを導入して、全体としてリスクは減ったけれど、そのリスクは一つのところに偏ってマイノリティのところに集積する可能性があるため、リスクが向かう先については気をつけた方がよい。

AI ロボット活用の格差・不平等の発生リスク【プロジェクト共通】

2050 年における社会において一定程度の AI ロボットが導入されていると想像すると、その AI ロボットを導入して活用できる環境を有する団体(企業、大学など)及び団体所属個人は比較優位性を持っていると考える。その

一方で、金銭的な面などから、そのような環境を有することが出来ない団体及び団体所属個人も存在することが想定される。そこで両者の間で利用に関しての格差が生じるリスクが考えられる。例えば、科学探究を行うための AI ロボットについて、企業においては資金力がある企業が優位性を確保するであろう。ただ、大学の研究機関においては、一部の AI ロボットの資産を他の大学に対して条件付きで共通的に利用できる環境を提供する可能性もあり、その場合は AI ロボット活用の格差は生じにくいと考える。

また、AI ロボットが社会の広い場面に導入されていくと、個人の生活環境にも影響を与えていくことが想像される。家庭内に AI ロボットが導入されていく可能性もあり、家庭の資産によって導入可否も変わってくるだろう。AI ロボットを導入できる資産のある家庭では、導入によるメリットを享受することができ、例えば余暇時間が増えるなど、生活の質が向上することが期待される。その一方で、導入資金が乏しい家庭では生活の質は変わらないため、その格差は広がるだろう。仮に、AI ロボットが共通のインフラとして、国や自治体の費用補助により多くの家庭に導入されたとした場合においても、その利用の仕方が分からない人も出てくるであろう。特に現在でもスマートフォンの利用に不慣れな高齢者が存在するように、2050 年の AI ロボットの利用に不慣れな人が生じる可能性はあるだろう。もしくは人が操作しなくても AI ロボットが適切な動作を行うような状況もあるかもしれない。

いずれにせよ、AI ロボットの導入によって、それぞれの場面(企業、大学、家庭など)の格差・不平等が生じることになるが、過去から現在においても形を変えて格差・不平等は存在してきたものであり、2050 年においてもその存在を完全に消し去ることは出来ないだろう。しかしながら、現状の格差・不平等を増大しない形で、可能ならば格差・不平等を縮小する形で、AI ロボットの導入が望まれることは言うまでもない。

人間に疑似する AI ロボットが作られた場合の不気味の谷の発生リスク【菅野プロジェクト】

目標3ではゲル状素材等を使った自己修復型の機構を開発しているプロジェクトもあり、指や手から開発している。将来、このような柔らかい材料で人間に見た目が近い人型ロボットが作られた場合に、そのロボットの手を人間が握手した場合に、ロボットの手の感触・温度次第で、不気味の谷¹⁵⁾の現象が発生する可能性があると考ええる。そもそも不気味の谷の発想は、森 政弘先生が 1970 年当時、義手の手と握手したときの不気味さを感じたことから始まっている。社会実装においては、不気味の谷を考慮した開発が必要と考える。

ただし、現時点の目標3のプロジェクトでは、見た目が人間に近いロボットを作ることは考えておらず、明らかに人間とは異なるロボットの指・手・腕を作ろうとしている。不気味の谷の現象は、見た目が人間に近いと起きる現象なので、現時点で想定しているロボットでは、この現象は起きないと考えている。

【本章のまとめ】

- AI ロボットの社会的な受容性について、雇用影響、事故発生、活用の格差の論点を考察した。
2050 年においては日本の労働力不足の問題を解決するために、AI ロボットの社会導入を受容せざるをえない状況にあるのではと推察する。

<社会的課題(2030 年及び 2050 年)>

- 雇用・労働への影響
- 事故発生に関する不安
- AI ロボット活用の格差・不平等の発生リスク
- 人間に疑似する AI ロボットが作られた場合の不気味の谷の発生リスク(2050 年)

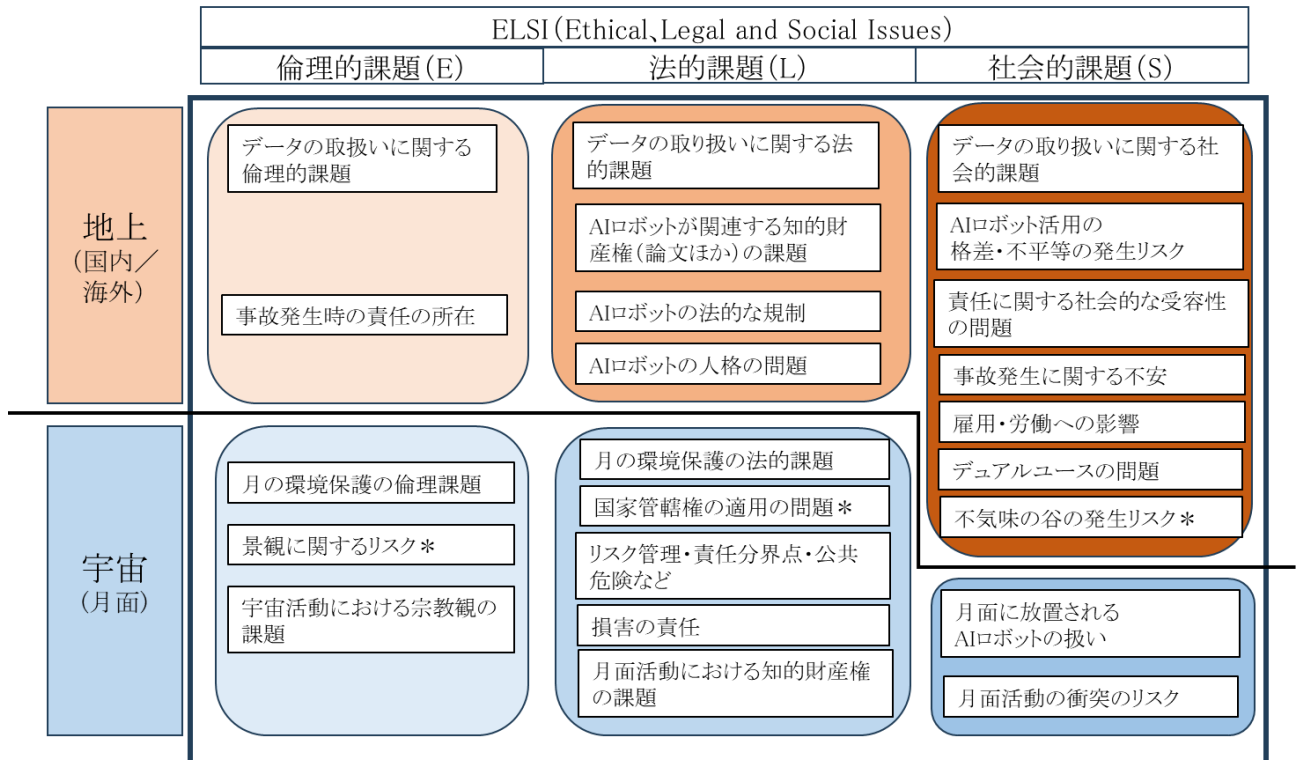
参考文献

- 1) 科学技術振興機構 研究開発戦略センター(CRDS),「科学技術・イノベーションの土壌づくりとしての ELSI/RRI 戦略的な科学技術ガバナンスの実現に向けて」,CRDS-FY2023-SP-01,2023 年 5 月.
- 2) European Commission, “AI Act”, <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>, (2025 年 11 月 12 日アクセス).
- 3) 大屋雄裕「AGIと社会と人類と」『情報通信政策研究』(総務省学術雑誌) 8 巻 1 号(2024 年),pp. 1A-19-30, https://www.jstage.jst.go.jp/article/jicp/8/1/8_19/_pdf/-char/ja, (2025 年 11 月 12 日アクセス).
- 4) 小塚壮一郎・佐藤雅彦編『宇宙ビジネスのための宇宙法入門[第 3 版]』(東京:有斐閣,2024),p.8.
- 5) 小塚壮一郎・佐藤雅彦編『宇宙ビジネスのための宇宙法入門[第 3 版]』(東京:有斐閣,2024),p.233.
- 6) 小塚壮一郎・佐藤雅彦編『宇宙ビジネスのための宇宙法入門[第 3 版]』(東京:有斐閣,2024),pp. 231-232.
- 7) 内閣府『令和 7 年版高齢社会白書(概要版)(PDF 版)第 1 章 高齢化の状況』, <https://www8.cao.go.jp/kourei/whitepaper/w-2025/gaiyou/pdf/1s1s2s.pdf>, (2025 年 11 月 12 日アクセス).
- 8) 国際連合広報センター「人口と開発」, https://www.unic.or.jp/activities/economic_social_development/social_development/population/, (2025 年 11 月 12 日アクセス).
- 9) 国立社会保障・人口問題研究所『人口統計資料集(2025)』,表 1-5,表 2-7, <https://www.ipss.go.jp/syoushika/tohkei/Popular/Popular2025.asp?chap=0>, (2025 年 11 月 12 日アクセス).
- 10) 国立社会保障・人口問題研究所『日本の将来推計人口(令和 5 年推計)』,p.123,p.147, https://www.ipss.go.jp/pp-zenkoku/j/zenkoku2023/pp2023_ReportALLc.pdf, (2025 年 11 月 12 日アクセス).
- 11) 国立社会保障・人口問題研究所『日本の世帯数の将来推計(全国推計)』,p.12, https://www.ipss.go.jp/pp-ajsetai/j/HPRJ2024/houkoku/hprj2024_houkoku.pdf, (2025 年 11 月 12 日アクセス).
- 12) リクルートワークス『未来予測 2040』,p.4, <https://www.works-i.com/research/report/item/forecast2040.pdf>, (2025 年 11 月 12 日アクセス).
- 13) 三菱総合研究所『生成 AI の普及が与える日本の電力需要への影響』2024 年, <https://www.mri.co.jp/knowledge/insight/policy/i5inlu000000np1p-att/nr20240828pec-1.pdf>, p.6,(2025 年 11 月 12 日アクセス).
- 14) 内閣府『令和 6 年度 年次経済財政報告』,第 2 章第 3 節, <https://www5.cao.go.jp/j-j/wp/wp-je24/h02-03.html>, (2025 年 11 月 12 日アクセス).
- 15) 日本ロボット学会,「日本のロボット研究の歩み,不気味の谷」, <https://robogaku.jp/history/integration/I-1970-1.html>, (2025 年 11 月 12 日アクセス).

付録

付録1. 目標3 AI ロボットの ELSI 論点整理の課題の俯瞰図

目標3 AI ロボットの ELSI 論点整理の課題の俯瞰図を示す。（＊印の課題は本文の表現を簡略化している。）



付録図1. 目標3 AI ロボットの ELSI 論点整理の課題の俯瞰図

ELSI 研究会関係者

(敬称略、有識者メンバーは五十音順、所属・肩書きは 2025 年 11 月時点のもの)

目標3 PD 福田 敏男 名古屋大学 名誉教授／名古屋大学 未来社会創造機構 客員教授

<ELSI 研究会メンバー>

統括責任者 友枝 敏雄 九州大学 名誉教授／大阪大学 名誉教授、目標3アドバイザー

有識者メンバー 青木 節子 慶應義塾大学大学院法務研究科 教授(令和7年3月31日まで)
千葉工業大学特別教授(令和7年4月から所属変更)

稲谷 龍彦 京都大学大学院法学研究科 教授(令和7年3月31日まで参加)

大屋 雄裕 慶應義塾大学法学部 法律学科 教授(令和7年4月から参加)

久木田 水生 名古屋大学大学院情報学研究科 准教授

寺田 麻佑 一橋大学ソーシャル・データサイエンス研究科 教授(令和7年4月から参加)

平野 孝典 桃山学院大学社会学部 准教授

各プロジェクトの会議参加者(延べ人数)

PM:プロジェクトマネージャー、PM 補:プロジェクトマネージャー補佐、PI:課題推進者、その他は研究参加者

菅野プロジェクト 菅野 重樹 PM 早稲田大学 理工学術院 教授

松本 治 PM 補 早稲田大学 次世代ロボット研究機構 上級研究員

小林 英津子 PI 東京大学 工学部 教授

桑名 健太 PI 東京電機大学工学部 教授

高橋 利枝 PI 早稲田大学 文学学術院 教授

板倉 陽一郎 早稲田大学 次世代ロボット研究機構 客員上級研究員

河島 茂生 早稲田大学 次世代ロボット研究機構 客員主任研究員(准教授相当)

河井 大介 早稲田大学 次世代ロボット研究機構 客員次席研究員

原島 大輔 早稲田大学 次世代ロボット研究機構 客員次席研究員

平田プロジェクト 平田 泰久 PM 東北大学 大学院工学研究科 ロボティクス専攻 教授

岡部 康平 PI 労働者健康安全機構 労働安全衛生総合研究所 機械システム安全研究グループ 上席研究員

稲邑 哲也 PI 玉川大学 脳科学研究所 教授

加藤 健治 PI 国立長寿医療研究センター 健康長寿支援ロボットセンター ロボット臨床評価研究室長

久保田 直行 PI 東京都立大学 システムデザイン学部 機械システム工学科 教授

下田プロジェクト 下田 真吾 PM 名古屋大学 大学院医学系研究科 特任教授

	志水 洋人	名古屋大学大学院医学系研究科 特任助教
原田プロジェクト	原田 香奈子 PM	東京大学 大学院医学系研究科／大学院工学系研究科 教授
	丸山 善宏 PI	名古屋大学 大学院情報学研究科 准教授
牛久プロジェクト	牛久 祥孝 PM	オムロンサイニックス株式会社 リサーチアドミニストレイティブディビ ジョン リサーチバイスプレジデント
永谷プロジェクト	永谷 圭司 PM	筑波大学 システム情報系 教授
	安川 裕介 PM 補	国際航業株式会社 事業統括本部国土保全部総合防災グループ 職員(リーダー職)
國井プロジェクト	國井 康晴 PM	中央大学 理工学部 教授
	藤吉 隆雄 PM 補	中央大学 研究支援室 PM 支援専門員 URA
	吉光 徹雄 PI	宇宙航空研究開発機構 宇宙科学研究所 教授
	宮口 幹太 PI	株式会社竹中工務店 技術研究所 未来・先端研究部 主席研究員
吉田プロジェクト	吉田 和哉 PM	東北大学 大学院工学研究科 教授
	小川 ちほ PM 補	東北大学 大学院工学研究科 技術補佐員
	山野辺 夏樹 PI	産業技術総合研究所 インダストリアル CPS 研究センター 上級主任研究 員
事務局 主担当	森 勇一	JST ムーンショット型研究開発事業部 技術主幹
事務局 副担当	久保 秀之	JST ムーンショット型研究開発事業部 技術主幹