

令和 4 年 6 月 16 日現在

機関番号：82646

研究種目：基盤研究(C) (一般)

研究期間：2018～2021

課題番号：18K01309

研究課題名(和文) 知的エージェントによる有害行為に対する刑事責任

研究課題名(英文) Criminal responsibility for harmful actions by intelligent agents

研究代表者

石井 徹哉 (Ishii, Tetsuya)

独立行政法人大学改革支援・学位授与機構・研究開発部・教授

研究者番号：20351869

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：本調査研究の当初の目的は、知的エージェントの製造者、提供者または管理者・使用者の刑事責任をどのような要件理論構成で肯定できるかということを示し、その関係において、知的エージェントそれ自体の主体性・責任についての考え方の方向性を検討するというものであった。最終的には、研究協力者の寄与等もあり、知的エージェントの主体性の哲学的基礎づけ方向性、正犯性の問題などさらなる発展的問題、さらには特別法上の行為主体性の検討も行うことができた。過失犯についても、製造物責任のみならず、過失要件との具体的なすり合わせ、自動運転自動車の実装に係る道路交通法条の問題を比較法的に研究した。

研究成果の学術的意義や社会的意義

いわゆるロボット又はAIについて、現行刑法における諸問題を網羅的に検討し、現状の解釈論を示すことができたほか、レベル4の自動運転自動車について、その社会実装のための道路交通法の日独比較法研究を先端的に実施し、今後の法解釈、運用の先行研究となり得るものとなった。

また、業法領域におけるAIの主体性と業法違反の問題についても、先行的研究となっている。

研究成果の概要(英文)：The initial purpose of this research was to clarify what kind of requirement theory structure could be used to affirm the criminal responsibility of the manufacturer, provider, or administrator/user of an intelligent agent, and to examine the direction of thinking about the subjectivity and responsibility of the intelligent agent itself in this context.

In the end, with the contribution of research collaborators, we were able to examine the philosophical basis for the subjectivity of intelligent agents, further developmental issues such as the problem of legitimate criminality, and the subjectivity of acts under special laws. Regarding negligence offenses, we were able to sort out not only product liability, but also the specific reconciliation with negligence requirements, examination of the issue of the article of the Road Traffic Law regarding the implementation of automated vehicles, and the concept of dilemma situations from the aspect of criminal law.

研究分野：刑事法

キーワード：AI ロボット 自動運転自動車 過失犯 行為主体 責任

1. 研究開始当初の背景

平成 29 年 6 月に閣議決定された『未来投資戦略 2017 —Society 5.0 の実現に向けた改革—』において、「ロボット新戦略」(平成 27 年日本再選戦略本部決定)を社会的実装の段階に移すことが示されている。「ロボット」の概念は、広範であるが、ここでは、「知的エージェント」による自動化が主として問題とされている。「知的エージェント」とは、特定の作業を機械的に実行するのではなく、独立で行動し、状況の変化を学習して適応する特質をもつ。そのために、大量のデータから高速に学習し、絶えず新たな問題解決規則に適応する機能が認められる。

このような知的エージェントにより社会活動の一部が自動化される場合、知的エージェントが人を介さず情報ネットワークを通じて情報をやりとりし、ある種自動的に動作することが想定される。しかし、これまでの刑事法制をみるならば、法制度それ自体が情報通信技術の進展に遅れをなし、弥縫策しか提供できず、体系性を欠如する状況にある(石井徹哉「保護客体としての情報」井田良ほか編『理論刑法学の探究』(2015 年)139 頁以下)。ましてや、知的エージェントによるネットワーク形成や人と知的エージェントとの混在する社会に対して、現行の法制度は、知的エージェントの自動処理に起因する事故に対してまったくなんの準備もできていない。

知的エージェントによる自動処理は、学習等により当初の想定外の状況に臨機応変になされることも可能となる。この場合、ある種の「人」としての判断がなされているとみることも可能である。自律的判断を下した「人」に第 1 次的な刑事責任を科するのが刑事法の本来の考え方である。この点と製造者の刑事責任の追及可能性を十分に精査されないでいるならば、伝統的な自由主義的な刑法もしくは責任刑法の根幹が揺るがされるにとどまらず、製造者への過度の刑事責任の追及によりロボティクスを中心とする科学技術の進展をも阻害しかねない。

さらに、従来、コンピュータソフトウェアについては、プログラム上の瑕疵が当然に存在するものと前提で、他者への使用が許諾されてきた経緯がある。プログラムの瑕疵により生じる損害は、少なくとも作成者に法的責任を追及するのではなく、瑕疵の改修により改善し、将来的な損害の回避への対応へと向けられてきた。しかし、知的エージェントの中核をなすコンピュータプログラムの設計上の瑕疵が直接人身的被害をもたらすものであり、従来のコンピュータプログラムの瑕疵の問題と同列に論じることは難しく、また、現行の刑法理論がどこまで通用し、あるいは妥当するのかは、それほど明らかではない。

2. 研究の目的

(1) 研究の全体像

本研究は、上記背景をふまえ、現行刑法及び現在の刑法理論が、知的エージェントが導入された社会において、どこまで機能し、知的エージェントがもたらす社会的に有害な事態に対処できるのかを明らかにすることにある。

先端的な科学技術に関する法規制については、ともすれば特別法の制定により対処し、または関係省庁による各種ガイドライン等で対処することがなされがちである。しかし、現行法の妥当領域を厳密に検討しないままの特別法の立法は、弥縫策にすぎない上、ガイドライン等による対応は、刑法の基本原則である罪刑法定主義を蔑ろにするものである。したがって、知的エージェントの導入により生じることが予想される問題について、まずは現行刑法の解釈による解決可能性を突き詰めることが必要となる。

また、知的エージェントを社会実装するにあたっては、その実装を想定する領域において、行政庁による各種規制が既に存在しており、これらの規制との整合性を図ることも必要となる。直近において社会実装されるうと予想される領域は、いわゆる自動運転自動車であるが、ここでは、特に道路交通法による規制との関係並びに事故時における人身傷害及び物損に対して誰がどのような範囲において刑事責任を追うのかを自動車運転処罰法、刑法上の過失致死傷罪、さらには道路交通法違反とについて検討することが必要であろう。

(2) 研究の個別課題について

本研究では、知的エージェントがもたらす社会的に有害な事態に対する刑事責任の在り方を明らかにすることを目的とするが、焦点は、知的エージェントの介在が刑法上の帰責にとってどのような意義を有するのかということにある。刑法が人に対する刑事責任を追求するためのものであることを前提にすると、知的エージェントの製造者、提供者または管理者・使用者(以下「製造者等」)の刑事責任をどのような要件、理論構成で肯定できるかということが問題となる。

情報システムと人とが複合する領域における刑事責任の所在について、これまできちんと検討されてこなかったといえる。本研究は、法倫理及び行政法学の知見をも踏まえ、状況の認知に基づき自動的に行動するロボットと人との共存する社会において、刑法がどのように機能しうるかという問題の手がかりをみいだすことを目標とし、ロボティクスの進

展を阻碍せずかつ安全な社会の実現のための刑事法のあり方を検討することができればと考えている。将来的にロボット一般の法倫理的課題を検討するための礎とすることを目的とする。また、ロボットの自律判断に対する刑事責任の法倫理的問題を検討の基礎とすることで、従来の刑法、刑事責任の構想の変革をももたらしうるとの波及効果も目的としている。

従来の過失犯論における解釈適用の可能性の検討

これを人としての行為者（製造者等）の側から見ると、知的エージェントのもたらした（人的）被害に対する過失犯の成否となる。ここでは、いまだ刑事責任では、十分に検討されていない製造物責任の問題を正面から取扱い、従来の刑事過失論との整合性を明らかにする必要がある。

知的エージェントそれ自体の刑事責任又は行為主体性の検討

問題を知的エージェントの側から見ると、知的エージェントの「主体性」の問題となる。ここで問題となる「主体性」は、犯罪能力や受刑能力という意味ではなく、社会的に有害な事態の帰責を知的エージェントの背後者に遡及し、その刑事責任を問いつるかという意味での機能的な概念である。この問題は、知的エージェントから攻撃を受ける者からすると、知的エージェントに対して対抗措置をとることが犯罪の違法性を阻却するかという問題としても展開される。この二つの面を軸として、知的エージェントの「主体性」を明らかにし、刑法上の帰責の根拠づけと限界づけを明らかにすることが必要となる。

派生的問題

本研究においては、上記の主たる問題の検討を基軸として刑法上の解釈論又は立法論を検討するものであるが、準備段階での研究状況及び本研究実施後の研究状況から、更なる派生的問題の解決が必要となることも想定している。

例えば、自動運転自動車の社会実装を想定する場合、道路交通に関する国際条約の動向、わが国の道路交通法の改正状況なども検討に入ることが予想されよう。そのほか、主体性との関係においては、知的エージェントそれ自体の刑法における位置づけを明らかにしたり、場合により共犯論への波及も検討することが想定される。専門職におけるアシストとして知的エージェントを導入する場合には、行為主体性の観点からいわゆる業法との整合性も視野に入ることとなりうる。

いずれにせよ、こうした派生的問題については、研究の進行状況に鑑み必要に応じて研究対象となりうるものと想定していた。

3. 研究の方法

本研究の実施にあたっては、研究代表者及び研究分担者のほか、博士後期課程に在籍する若手研究者を研究協力者とし、またドイツにおけるロボットと刑法に関する研究の第一人者であるエリック・ヒルゲンドルフ教授（ヴェルツブルク大学）を研究協力者として研究組織を構成した。なお、研究年度の進展に伴い、科研費申請資格を得た場合は、若手研究者を順次研究分担者とし、研究に必要な予算執行の自由度を高めた。

研究の実施にあたっては、研究代表者である石井が全体的な統括を図るとともに、主として行為主体性及び刑事責任論について調査研究を実施し、研究分担者である若手研究者の研究協力者については、本研究課題に関して各自の専門との関連づけて関心のあるテーマを配分し、比較法研究を中心に調査研究することとした。また、比較法研究にあたっては、研究協力者のヒルゲンドルフ教授と連携を図ることとした。

また、研究初年度に研究分担者が渡独して、本研究に関する人脈を形成し、ドイツにおけるより幅の広い情報収集を可能とするように努めた。

具体的な研究の実施の仕方としては、研究代表者、研究分担者及び研究協力者が必要とする文献を収集し、日独における関係研究のサーベイを行い、それをもとに各自のテーマについて研究を進めることとし、定期的に研究会を開催して、研究状況及び研究成果の報告並びにこれらについての議論を行うことで研究を進めた。当初は、ドイツにおけるAIまたはロボットと刑法に関する文献の購読、紹介から始め年度を重ねることで、各自の具体的な研究成果の検討へと進めるものとなった。

最終的には、各自が本研究に関して論文を執筆し、それをもとに議論を行ったほか、本研究の成果として書籍を発行することとし、それに向けて各自が論文を執筆することにし、各自の原稿を元に研究会において議論を重ねた。

なお、COVID-19の影響により、全体として研究遂行及び研究の打ち合わせが困難になり、研究が停滞したため、研究遂行による成果を出すという観点から研究期間を延長している。

4. 研究成果

- (1) 知的エージェントといっても、その適用領域は多様であり、これを抽象的、一般的に論じても有益ではないと判断し、まずは特定の機能を実現するために構築され、社会実装が目指されているものとして、いわゆる自動運転自動車を念頭に置きつつ、全体の研究を開始した。

そのなかで、まず問題となったのは、自動運転自動車におけるいわゆるトロッコ問題、人身障害発生という緊急状況の回避に伴う他の人身傷害発生可能性への対応である。こ

の点に関して、具体的な理論構成については異なるものの、結論としては、現実問題として極めて希なケースであり、これを基礎として全体の理論構成を構築するのは困難ではないかということとなった。そのため、社会実装におけるプログラミング等のシステム構成に関する倫理的問題とするか、そもそも社会実装にあたり一定のリスクが内在することはやむをえず、当初のリスクの範囲内であれば、製造者も免責される構成が望ましいという方向となった。

製造者の責任については、潜在的なバグ等については、プログラミングの性質上これを殊更に取り上げて過失責任を追求するのは、困難であり、むしろプログラミングの脆弱性又は不具合が確認できた場合に、速やかにこれに対応することについて過失責任を問題とするのが適切ではないかということへと至った。

- (2) 行為主体性に関するものとして、犯罪能力及び受刑能力、知的エージェントが介在する場合の結果帰属及び共同正犯論、知的エージェントの欠陥による人身損害が発生した場合の過失責任の所在及び理論構成、緊急状況下におけるインテリジェントエージェントの動作に関する問題、ウィーン条約のもととジュネーブ条約のもとでの道路交通法における規制のあり方の比較を各自分担して、調査研究を実施した。
行為主体性のうち、犯罪能力及び受刑能力については、法哲学的知見を基礎として責任の本質、自由意思の問題を検討し、様々なモデルを検討するなかで、知的エージェントに対する刑事責任を基礎づけるモデルがありうることを明らかにした。ただ、それが刑法理論としてなお受入可能といえるかは、さらなる検討が必要である。
共犯論における知的エージェントの問題は、知的エージェントの行為主体性と関係し、現状においてはなお道具性にとどまっているのではないかと結論にとどまっている。
いわゆる業法上の主体性についても、知的エージェントがその結論を出す過程が不透明であるとはいえ、その判断をもって直ちに専門的判断とすることは困難であり、専門性をもつ人間の補助にとどまるものであり、また、この限りでいわゆる AI の判断をもとになんらかの業務行為を行い、損害が発生した場合は、AI の判断を利用した行為者の過失責任を追求すべきものとするに至った。
- (3) 道路交通法に関係する部分については、ウィーン条約及びジュネーブ条約、さらにはドイツにおけるレベル4を想定した自動運転自動車に関する法改正までフォローすることができ、わが国における同様の道路交通法改正と比較可能なものとすることができた。
- (4) 研究成果のとりまとめについては、研究協力者を含めて書籍として公刊することとしている。本研究課題に関して、ロボットまたは AI に関する日独の議論状況の概要とその比較検討、AI それ自体の行為主体性、AI による社会的損害に対する刑事責任のあり方、自動運転自動車等に関する道路交通法上の課題を内容としている。については、一般的な行為主体性、処罰可能性のみならず、AI またはロボットと人間との競合によるある種の「共犯」的責任まで検討している。ヒルゲンドルフ教授にも本研究協力者として研究とりまとめへの参画してもらっている。最終的に、研究期間未までに全員の原稿が完成しており、現在、出版社に依頼して公刊準備を進めている。
- (5) 本調査研究の当初の目的は、現行刑法及び現在の刑法理論が、知的エージェントが導入された社会において、どこまで機能し、知的エージェントがもたらす社会的に有害な事態に対処できるのかを明らかにすることとしていた。具体的には、知的エージェントの製造者、提供者または管理者・使用者の刑事責任をどのような要件理論構成で肯定できるかということを中心にし、その関係において、知的エージェントそれ自体の主体性・責任についての考え方の方向性を検討するというものであった。しかしながら、最終的には、研究協力者の寄与等もあり、知的エージェントの主体性の哲学的基礎づけ方向性、正犯性の問題などさらなる発展的問題、さらには特別法上の行為主体性の検討も行うことができた。過失犯についても、製造物責任のみならず、過失要件との具体的なすり合わせ、自動運転自動車の実装に係る道路交通法条の問題の検討、ジレンマ状況の考え方についても、刑法の側面から整理することができた。

5. 主な発表論文等

〔雑誌論文〕 計6件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 石井 徹哉	4. 巻 56-2
2. 論文標題 AIに関する刑法的課題	5. 発行年 2019年
3. 雑誌名 罪と罰	6. 最初と最後の頁 5-19
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 岡部雅人	4. 巻 下
2. 論文標題 「許された危険」とは何か その序論的考察	5. 発行年 2021年
3. 雑誌名 『寺崎嘉博先生古稀祝賀論文集 [下巻] 』	6. 最初と最後の頁 71-88
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 山下裕樹	4. 巻 45
2. 論文標題 AI・ロボットによる事故の責任の所在について - 自動運転車の事案を中心に -	5. 発行年 2019年
3. 雑誌名 ノモス	6. 最初と最後の頁 95-105
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 エリック・ヒルゲンドルフ（富川雅満訳）	4. 巻 50
2. 論文標題 自動運転と法	5. 発行年 2018年
3. 雑誌名 龍谷法学	6. 最初と最後の頁 597-620
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 根津洸希	4. 巻 50
2. 論文標題 AI 技術を巡る刑法的問題の概説と解決の試み (部分的) 自動運転技術を一例に	5. 発行年 2021年
3. 雑誌名 中央大学大学院研究年報	6. 最初と最後の頁 85-100
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Koki NEZU	4. 巻 2019-4
2. 論文標題 Strafrechtlicher Problemaufriss von (teil)autonomen Fahrzeugen in der Gegenwart und Zukunft	5. 発行年 2020年
3. 雑誌名 Hanover Law Review	6. 最初と最後の頁 268-274
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計1件

1. 著者名 石井徹哉編著	4. 発行年 2022年
2. 出版社 成文堂	5. 総ページ数 400
3. 書名 ロボットと刑法	

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	岡部 雅人 (Okabe Masato) (30453989)	国土館大学・法学部・教授 (32616)	

6. 研究組織（つづき）

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	伊藤 嘉亮 (Ito Yoshisuke) (00837792)	早稲田大学・社会科学総合学院(先端社会科学研究所)・助教 (32689)	
研究分担者	富川 雅満 (Tomikawa Masamitsu) (80781103)	九州大学・法学研究院・准教授 (17102)	
研究分担者	山下 裕樹 (Yamashita Hiroki) (20817150)	神戸学院大学・法学部・講師 (34509)	

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究協力者	根津 洸希 (Nezu Koki)		
研究協力者	田村 翔 (Tamura Sho)		
研究協力者	今井 康介 (Imai Kosuke)		
研究協力者	ヒルゲンドルフ エリック (Eric Hilgendorf)		

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------