博 士 論 文

AIの利活用における刑法上の諸問題
―利用者と製造者の刑事責任を中心に―

(The Criminal Law Problems in the Use of Artificial Intelligence
- Focus on the Criminal Liability of Users and Producers -)

2023年3月

立命館大学大学院法学研究科

法学専攻博士課程後期課程

日原　拓哉

立命館大学審査博士論文

AIの利活用における刑法上の諸問題
―利用者と製造者の刑事責任を中心に―
(The Criminal Law Problems in the Use of Artificial Intelligence - Focus on the Criminal Liability of Users and Producers -)

2023年3月

March 2023

立命館大学大学院法学研究科

法学専攻博士課程後期課程

Doctoral Program in Law

Graduate School of Law

Ritsumeikan University

日原　拓哉

HIHARA Takuya

研究指導教員：　安達　光治 教授
Supervisor： Professor ADACHI Koji

目次

細目次

はじめに

深層学習がその火付け役となり、21 世紀初頭から現在に至るまで継続する第 3 次 AI ブームも佳境に差し掛かる。近時、自動運転車、産業用ロボット、介護用ロボットなど、社会生活上においてますます AI・ロボットが浸透してきている。日常生活を円滑にするために、これらに判断を委ねることが徐々に増えつつある。例えば、自動運転車、家事や手術室、軍が運営する監視室で用いられるロボットだけでなく、車の駐車や予定のリマインダーを設定することなどが挙げられよう。しかし、それに伴う人的損害も存在し、例えば、股関節手術にミリングロボットを使用したところ神経や筋肉に損傷を与えた事例[1]、車線維持支援システムが原因で、母子が事故死した事例[2]、バウナタールにある VolksWagen 工場で作業員がロボットによって死亡した事例[3]がある。さらに、将来的に想定されうるものとして市場経済を脅かす事例も考えられる[4]。具体的には、AI・アルゴリズムを利用した投資によって相場操縦に該当する取引が行われる事例未公開情報をデータベースに記録し、AI・アルゴリズムにインサイダー取引をさせる事例等が考慮される。

AI の利活用をめぐっては、とりわけ、自動運転車の関与する交通事故事例が注目されてきた。例えば、Tesla 社製の自動運転車（レベル 2）の「オートパイロット」を搭載した自動運転車を運転中、眠っていた疑いがあるとして過失致死罪で訴追した事例[5]や、Uber 社製の自動運転車（レベル 4）の公道実験中の事故（2018 年）に関し、2020 年 8 月 27 日、同車に乗車していたドライバーが過失致死罪で起訴された事例[6]が挙げられる。後者においては、同人の裁判について 2021 年 8 月 10 日（現地時間）に第一審が開廷される予定であったが延期されている[7]。このように、AI の利活用に関する利用者もしくは製造者の刑法上の問題

---

[1] *Caetano da Rosa*, Robodoc - Zukunftsvisionen und Risiken robotisierter Spitzentechnik AI im Operationssaal, Technikgeschichte 74 (2007), S. 291 ff.

[2] 部分的自動運転の乗用車のドライバーが脳卒中になり,ハンドルを握れなくなった事例である。車線維持支援システムが車両を安定させ、意識を失ったドライバーの乗った車両を高速で町中に誘導し、そこで衝突した。システムの介入がなければ、車両はすでに町の入り口で停止していたとされる。*Hilgendorf*, Automatisiertes Fahren und Strafrecht - der Aschaffenburger Fall", DRiZ 2018, S. 66 を参照のこと。

[3] FAZ v. 01.07.2015 (Roboter tötet Arbeiter bei VW in Baunatal").Volks Wagen 社の工場で勤務する 22 歳の男性は、電気モーターの新規生産ラインでロボットのセットアップに追われていたところ、ロボットにつかまり、金属板に押し付けられ、胸部に重度の打撲傷を負った。一命をとりとめたものの、その後、病院で亡くなったという。

[4] アルゴリズム・AI の利用を巡る法律問題研究会「投資判断におけるアルゴリズム・AI の利用と法的責任」金融法務（2019 年）を参照。

[5] *Hawkins,* (2020, September 18). Tesla owner in Canada charged with ‘sleeping’ while driving over 90 mph. THE VERGE. (https://www.theverge.com/2020/9/18/21445168/tesla-driver-sleeping-police-charged-canada-autopilot)（最終アクセス 2022 年 11 月 27 日）

[6] Uber's self-driving operator charged over fatal crash. (2020, September 16). BBC NEWS. https://www.bbc.com/news/technology-54175359 (最終アクセス 2022 年 11 月 27 日)

[7] *Stern,* (2021, May 12). Trial Delayed for Backup Driver in Fatal Crash of Uber Autonomous Vehicle. PHOENIX New Times. https://www.phoenixnewtimes.com/news/uber-crash-arizona-vasquez-herzberg-trial-negligent-homicide-

は、主に「自動運転車」の事故事例を引き合いに、2017 年ごろより検討されてきたが[8]、裁判例などの実務的見解が存在しないまま議論されてきた。

これに対し、Tesla 社製のレベル 2 自動運転車（Model X）による東名高速道路での死傷事故（横浜地判令和 2 年 3 月 30 日 判例秘書 LLI/DB L07550489）のように、実際に AI を利用した製品に起因する判例が実際に現れるようになった。そのため現在は、既存の議論と実務的見解との妥当性について再検討すべきものと思われる。

自動運転車を基調とした既存の議論では、主に SAE 規準でのレベル 1~2／3／4~5 に段階分けをした検討がなされてきた。そこでの論調として、レベル 1 ないしは 2 においては利用者も製造者もレベル 0（普通自動車）と同一の注意義務が課されるものとされ、レベル 3

---

charge-11553424.(最終アクセス 2022 年 11 月 27 日)

[8] ドイツにおいては、いわゆる RobotRecht チームによる研究が知られている。*Hilgendorf/Beck* (Hrsg.), Jenseits von der Maschine, Robotik und Recht 1 , Nomos, 2012 を皮切りに、主に AI の利活用に関する複数領域（例えば、自動運転車の民事法的問題・刑法上の問題やジレンマ問題、データ保護、AI の権利主体性など）の問題を扱う研究書が多数刊行されている。2022 年 11 月現在では、その研究書も 28 巻まで刊行されており依然としてその研究は盛んである。

また日本でも、それに続く形で自動運転車をめぐる法的課題を中心に多数の論稿が上梓されている。ここではその全てをあげることはできないが、以下の文献が挙げられる。弥永真夫・宍戸常寿編『ロボット・AI と法』深町晋也「ロボット・AI と刑事責任」（有斐閣、2018 年）209 頁以下、根津洸希「ロボット・AI に対して『刑罰』を科すことは可能か」法学新報 125 巻 11 号（伊藤康一郎先生追悼論文集、2019 年）475 頁以下、川口浩一「ロボットの刑事責任 2.0」刑事法ジャーナル 57 号（2018 年）4 頁以下、今井猛嘉「自動車の自動運転と刑事実体法―その序論的考察」西田典之先生献呈論文集（有斐閣、2017 年）519 頁以下、佐久間修「AI と刑法・序説」名古屋学院大学論集社会科学編 55 巻 1 号（2018 年）107 頁以下、遠藤聡太「人工知能(AI)搭載機器の安全性確保義務と社会的便益の考慮」法律時報 91 巻 4 号 19 頁以下、今井猛嘉「AI 時代の刑事司法」罪と罰 222 号（2019 年）、根津洸希「ロボット AI に対して『刑罰』を科すことは可能か」法学新報（中央大）125 巻 11 号（2019 年）475 頁以下、坂下陽輔「人工知能の開発・利用における過失―自動運転車と過失を題材に」法律時報 91 巻 4 号（2019 年）13 頁以下、笹倉宏紀「人工知能の法規制における行政手続と刑事手続」法律時報 91 巻 4 号（2019 年）41 頁以下、石井徹哉「AI に関する刑法上の課題」罪と罰 222 号（2019 年）5 頁以下、稲谷龍彦「人工知能搭載機器に関する新たな刑事法規制について」法律時報 91 巻 4 号（2019 年）54 頁以下、稲谷龍彦「ロボット事故の刑事責任」日本ロボット学会誌 1 巻 38 号（2020 年）37 頁以下。

またドイツ法との比較を扱ったものとして、根津洸希「ロボットの処罰可能性を巡る議論の現状について」比較法雑誌 51 巻第 2 号（2018 年）145 頁以下、伊藤嘉亮「エリック・ヒルゲンドルフ『ロボットは有責に行為することができるか？ 規範的な基本語彙の機械への転用可能性について』（文献紹介『ロボットと法』シリーズの論文紹介(1)）」千葉大学法学論集 31 巻 2 号（2016 年）136 頁以下、田村翔「サシャ・ツィーマン『機械の本性とは何であったか？ 機械刑法をめぐる議論について』（文献紹介『ロボットと法』シリーズの論文紹介(2)）」千葉大学法学論集 31 巻 3・4 号（2016 年）87 頁以下、根津洸希「スザンネ・ベック『インテリジェント・エージェントと刑法 過失、答責分配、電子的人格』（文献紹介『ロボットと法』シリーズの論文紹介(2)）」千葉大学法学論集 31 巻 3・4 号（2016 年）105 頁以下。米国の議論を扱ったものとして、松尾剛行「自動運転車と刑事責任に関する考察 ロボット法を見据えて」早稲田大学大学院法務研究科臨床法学研究会 Law and practice 11 巻（2017 年）73 頁以下など。

では、当該自動車の利用者たる「運転者」は、普通自動車のそれに加えて、オートパイロット時におけるオーバーライドに対応できるように運転する義務(道交法71条の4の2参照)や点検義務・整備義務（「自動運行装置」の利用者・運行供用者の義務）が加わる。また、製造者に関しては明確な言及はなされていないものの、普通自動車のそれと同一と結論づけるものが多い。レベル 4 ないしは 5 では、利用者については道交法上の「運転者」概念から外れるため、一見すると交通事故事例において利用者の道交法上の義務やそれに基づく過失責任を問うことはできないように見える[9]。それに対して製造者に対しては、普通自動車のそれと同一のスキームが妥当するという見解もある[10]。

AI（ここでは自動運転車）の利用によって人に死亡結果や傷害結果が生じた場合、AI をめぐる主体―利用者、販売者、開発製造者、プログラマー―にどのように刑事責任が帰属されるべきであるのかが問題となる。この点につき、従前の製造物による法益侵害結果惹起との注意すべき相違は、例えばAIを搭載した自動運転車（レベル4以上）が死傷事故を起こした場合、その事故原因にかかる機序がブラックボックス化しうることにある。このことは、責任所在の証明を困難にする可能性がある。すなわち、製造者ないしは利用者が当該義務を履行していれば結果が発生しなかったということがいえず、客観的に当該結果が帰属される主体が存在しない可能性、すなわち「帰属の間隙」が存在するおそれがある。もっとも、そのようなリスクをはらむことをあらかじめ上記主体に加えて社会全体が許容しており、それがAIを搭載した自動運転車を利用することに対して通常有すべきリスクであるとされるならばこのような問題は生じない。しかし、そもそも自動運転車（SAE基準でのレベル4以上）は実用化されておらず、現状はそのような状況にあるかは不明である。もっとも、その自動運転車が普通自動車の有するリスク、すなわち普通自動車による交通事故件数よりも少なくなるというのであれば、自動運転車の有するリスクを社会が許容しうる可能性は指摘される[11]。むろん、このことは現状では不明確なので、なおのこと現行法の解釈によって、誰にどのように答責を帰属させるのか、ないしは、帰属させることはできないのかを仔細に検討する必要がある。

考えられる帰属のモデルとして、例えばいわゆる「連帯責任」のモデルがある。これは、当該結果の発生に関与した主体に結果を分散的に帰属させる着想であり、このような分散的な結果帰属により、死傷結果に対する「帰属の間隙」を回避することはできるといえる。しかし、このモデルでは、とりわけ開発製造者側にとって、将来的な発生結果が自己に帰属されるリスクを背負ってまで製品を開発することを意味する。それは、製造開発者のモチベ

---

[9] もっとも2022年道路交通法改正により、いわゆるレベル4の自動運転車の利用者は「特定自動運行実施者」と定義され、その者に対する義務も制度化された。この改正については、樋笠尭士「自動運転レベル4における刑事実務―道路交通法改正案の分析と提案―」捜査研究858号（2022年）25頁以下が詳しい。

[10] 例えば、山下祐樹「AI・ロボットによる事故の責任の所在について」ノモス45巻(2019年)108頁。

[11] 松宮孝明「自動運転をめぐる刑事法的諸問題」立命館法学395号（2021年）4頁以下。

ーションという点からみて、開発を阻害する要因となり、結果として科学技術の進展を阻むことにもなりうる。利用者側についても、果たして、自己や第三者に生じた損害につき、過失責任が帰属されるリスクを背負ってまでAIを利活用するのか、という疑問が残る。それゆえ、損害結果の発生につき、分散的に答責を帰属させるという構想に立つ場合[12]、却って市民社会における市民の平穏な生活を法が脅かすおそれがあり、科学の発展を阻害するおそれがあるのではないかという懸念も存在する[13]。

解決手法と一つとして、AIの利活用における刑法上の答責帰属を議論する文脈で、AIそのものに刑罰を科すことができれば、このような問題は生じないとする見解もある[14]。しかし、AIそのものに刑罰を科すとする考え方については、因果的に結果を惹起せしめた自然人を特定する努力を放棄することにつながるのではないかという懸念が指摘される[15]。そのため、この見解を否定しつつ、AIの利活用に関与する人間の主体のうち、誰に当該結果について帰属されるのかを検討する必要がある。すなわち大別して、製造者と、管理者含む利用者それぞれの立場に応じ、従前の製造物とは異なり、予測不可能な判断をなしうるAI製品を利用し、これによって生命や身体といった法益が侵害された場合に、どのようにして刑法上の解決を図るべきかを検討する必要がある。

この背景には、AI製品によって利用者ないしは第三者の生命・身体が侵害された場合、製造者の刑法上の責任（刑事製造物責任）の前提となる、「必要な注意義務」の内容は、AIを搭載していない製品のそれと同一である否かの検討を改めて行うべきではないかという問題意識がある。そのため、改めて製造者に課せられる注意義務の内容を確定させる必要がある。

また、少なくともAIによる動作が介在する以上、その利用者には、それが搭載されていない製品以上の注意が課されるのではないか、すなわち利用者はAI製品をどの程度まで「信頼」することが許されるかということを考慮しなければならない。例えば、Tesla社製自動運転車の事故事例においては、異常な走行動作の原因が、センサーの欠陥ないしは限界なの

---

12 これを問題にするのは *Beck*, Die Diffusion strafrechtlicher Verantwortlichkeit durch Digitalisierung und Lernende Systeme, ZIS 03/2020, S.41 ff.など。

13 例えば、経済産業省「GOVERNANCE INNOVATION Ver.2: アジャイル・ガバナンスのデザインと実装に向けて」（2021年）66頁（https://www.meti.go.jp/press/2020/07/20200713001/20200713001-1.pdf :2022/05/29閲覧）など。

14 松宮孝明「自動運転と法」学術の動向（2020年5月）59頁以下、松宮・前掲（注10）では以下に示す論点の検討を簡単に批評する。ただし、それを可能にしようとすると、伝統的な刑罰理論にどこまで適合するのかを以下のプロセスで検討する必要性があるとされる。すなわち、①AIを人間と同レベルの権利主体として見なしてよいのか、②AIを処罰することに意義はあるか、③AIの処罰として考えられるものは「刑罰」であるのかというものである。なお、中国におけるAIと刑事責任の議論については、劉憲権（孫文訳,松宮孝明監訳）「人工知能時代における刑事責任の変遷」立命館法学396号（2021年）467(969)頁を参照。そこではAIに対する刑罰が肯定的に描かれている（486頁以下）。

15 この指摘は、山下・前掲（注10）106頁にもみられる。

かが判然としていない。そうすると、このような事故が発生しないように、利用者に対しては当該「AI の誤動作に対応できるように利用する」という新たな義務が課されることになろう。もっとも、このことは AI 製品の利用促進を阻害する要因にもなりうるのではないかという疑念、およびこれは利用者のみならずその背後に存在する保有者・運行供与者にも及ぶのではないかという問題もある。

以上のような問題意識のもと、本稿では AI 製品に由来する事故に関する刑事責任の検討を行う。そのような事例として、先に取り上げてきた自動運転車の事例のように、人間の生命・身体を侵害するものが想定されるが、それればかりでなく、近時その利活用が注目される投資アルゴリズム AI やソフトウェア・エージェントのような AI 製品は人間の財産を侵害することがありうる。具体的には、実体を持たないプログラムベースの AI の利活用によって刑法上の結果をもたらす事例が存在する。例えば、AI・アルゴリズムを利用した投資システムを利用したところ、利用者の認識なく、AI の判断により証券犯罪としての相場操縦取引やインサイダー取引が結果として行われてしまう事例や、さらに、ネットワーク化された経営判断に供される AI が協働してカルテルを締結する事例などについて、これらの経済犯罪には過失を処罰する規定が存在しないので、利用者の可罰性を肯定することは難しい。なぜなら、そもそも、不公正取引行為の過失処罰規定が存在しない以上、不公正取引をしているという認識のない利用者処罰を考慮すること自体に疑念はあるからである。これらの問題については、自動運転車等の製品における解釈とは別に、立法上の提言も含めて検討を行う必要がある。

最後に、本検討ではあくまで現状の技術水準における AI 製品の利活用に関するものが主となるところ、将来的には AI を搭載した自律的兵器のように、人間社会を脅かすような利用形態に対する危惧も示されている。このような AI 利用に関する研究を事前に抑制する刑法上の規制の可能性をも考察すべきあろう。

その際、まずは AI に関連する用語法の定義及び実例を確認し、AI に関する概念の明確化および本論文で想定する AI を示す（第 1 章)。さらに、現状の AI 製品および、将来的な AI 製品を基調とした刑法上の問題の検討を、その解決手法に関する議論を敷衍しつつ、最終的には利用者側の行為についての検討、そして製造者側の行為についての検討を刑事製造物責任の文脈に即して検討を行う（第 2 章)。さらに、人間の財産を侵害する特別な問題として、AI を通じた経済犯罪、そして AI が行為客体となるコンピュータ犯罪についての問題を検討し（第 3 章)、最後に今後の AI 開発における規制を刑法上の観点から検討するものとする（第 4 章)。

第１章　AI 概念の明確化

AI 製品の利活用をめぐる刑法上の諸問題に関する先行研究においては、AI の深層学習（ディープ・ラーニング）が因果関係の機序の不明確化、それに伴う予見可能性認定の困難性をもたらすとされるが、その前提である AI 概念や深層学習の定義理解が論者によって異なり、その上で予見可能性認定の検討を進められることが多いようである[16]。この点、概念を明確にしなければ正確に検討を進めることができない。このような問題意識のもと、本論文での検討にあたって、まずは AI やその周辺概念である機械学習や深層学習などの概念を AI 研究の歴史から遡る形で明らかにし、その上で想定すべき AI を措定する。

第１節　AI の歴史

人工知能という言葉は、1955 年 8 月 31 日に開催されたダートマス会議における、*John McCarthy, Marvin Minsky, Nathan Rochester,Claude Shannon*,"proposal for the Dartmouth Summer Research Project on Artificial Intelligence"という宣言書の中で初めて姿を表した[17]。当時ダートマス大学の数学の助教授であった McCarthy はこの宣言書で、「言語を使い、抽象化し、構想を展開し、目下人間のみにしかなしうることのできない種類の問題を解決し、自己改善を続けることができる機械をいかに作ることができるか」[18]という目標を掲げた。ここには、「推論」と「探索」を基軸に置く AI のみならず、ニューラルネットワークや人間の言葉のコンピュータ処理することも含まれていた。

この McCarthy らの宣言は、当時の AI 領域の研究において時代を先行していたという事実のみならず、ダートマス会議において AI が楽観的に理解されていたことも窺える。このようないわばユートピア的熱狂に駆り立てられる形で、McCarthy らは、専門家集団が数ヵ月間集中的に取り組めば、個々の研究分野での飛躍的進歩が達成されるとも述べた[19]。もっとも、このダートマス会議において飛躍的進歩をもたらしたというわけではないとされる[20]。

---

16 この指摘は、石井徹哉「AI に関する刑法上の課題」罪と罰 222 号（2019 年）6 頁にもある。例えば、今井・前掲（注 2）罪と罰 222 号（2019 年）25 頁は AI を「データ処理を超高度化させる」ものだとし、稲谷・前掲日本ロボット学会誌 1 巻 38 号（2020 年）37 頁以下は深層学習を「大量のデータを統計的に処理することで一定の法則性を見出し、自らそれに従って振舞う」ものとする(AI の定義はない)。また、例えば、佐久間修「AI の刑事責任―否定説の観点から」刑法雑誌 59 巻 2 号 159(297)頁以下では、AI 研究の展開を示しつつ、AI を「明文化しやすい行動原則にもとづくルールベースの AI」と、「統計 ・確率型の AI であって、『隠れマルコフモデル』とか『ベイジアンネットワーク』などと呼ばれる」AI と、「人間の脳機能を参考にしたニューラルネットワークであり、ディープラーニングによる各種のパターン認識をおこなう」AI の組み合わせであるとするが、深層学習については「多層的ニューラルネットワークによる情報処理が可能」なものであるとして、その効果のみが説明されるにすぎない。このように、論者によって AI や「深層学習」に対する理解には濃淡があることがうかがえる。

17 *Russel/Norvig,* Artificial Intelligence –A modern approach-, 4th ed, 2021, p.18.

18 *McCarthy, Minsky, Rochester, Shannon*, A proposal for the Dartmouth summer research project on artificial intelligence (Aug. 31st 1955), p.2.

19 *Ibid*.

20 *Russell/Norvig, supra* (fn.17), p.18

とはいえ、人工知能という言葉が初めて定着し、それ以降、当時の研究をはるかに超えたセンセーションを巻き起こしたことは否定し得ない。これが 1960 年代に始まる第 1 次 AI ブームである[21]。

ダートマス会議後も AI 研究はさらに進められた。例えば、Sammuel は、常に自分自身と対戦することで、徐々にそのゲーム戦略上の実力を向上させるようなコンピュータ・チェッカーを開発し、最終的にこのプログラムはチェッカー上級者のレベルに至ったという[22]。また、Newell と Simon は、Principia Mathematica における多くの数学上の定理の証明を可能にする記号処理プログラムである Logic Theorist を開発した[23]。むろん、John McCarthy 自身も、自らを改善するプログラムを書くことを可能にするプログラミング言語である「LISP」の開発者[24]としては無視できない。しかし、これらの AI 分野の先駆的開発には決定的な問題点があった。それは、どのプログラムもトイ・プロブレムを主に扱っているため、適用領域が非常に限定的であったことである。例えば、米軍の研究機関である DARPA は、冷戦時代にロシア語の文章を英語に自動翻訳するプログラムの開発に着手したものの、その開発は難航し、結果として研究資金が大幅に削減されるということがあった[25]。

AI が再び脚光を浴びたのは、1980 年代であった（第 2 次 AI ブーム）。第 1 次 AI ブームとは対照的に、ここでは「知識」を用いたエキスパート・システムがその中心に置かれている。この背景にあるのは、実世界に対応するシステムを開発するにはシステムが現実世界における膨大な知識を有する必要があることが強く認識されたことである[26]。これに対応する形で、ある専門分野の知識を取り込み、推論を行うことで、まるでその分野のエキスパートであるかのように振舞うプログラムの開発が進められた。その一例として、スタンフォード大学によって開発された MYCIN が挙げられる。これは、伝染病の血液疾患患者に対して質問に回答させる形で診断を行い、感染した細菌を特定し、そのうえで適切な抗生物質を処方するように設計されていた[27]。しかし、このシステムにも限界があり、莫大な知識を与えることによって、論理的矛盾や一貫性の喪失が生じた。とりわけ、「けだるさ」や「痛い」などの表現については個別に定義付けをしなければならなかった。このように、知識を与えることの困難性が露呈した結果、第 2 次 AI ブームはその終焉を迎えた。

その後しばらくは AI ブームに裏打ちされた AI 研究は下火であったが、それを覆す決定的な出来事があった。それは、1997 年の IBM によるチェスプログラム Deep Blue が、当時のチェス世界チャンピオンであった Kasparov に 3.5：2.5 で勝利を収めたことである[28]。こ

---

21 松尾豊『人工知能は人間を超えるか―ディープラーニングの先にあるもの』（角川 EPUB 選書、2015 年）81 頁。

22 See *Scheffer*, One Jump Ahead: Challeging Human Spremacy in Checkers, Springer, 2009, chap.6.

23 *Russell/Norvig, supra* (fn.17), p.18

24 *Russell/Norvig, supra* (fn.17), p.19.

25 *Russell/Norvig, supra* (fn.17), p.30.

26 小林一郎『人工知能の基礎』（サイエンス社、2008 年）4 頁。

27 松尾・前掲（注 21）87 頁。

28 *Russell/Norvig, supra* (fn.17), p. 27.

のコンピュータ戦で注目すべきは、Kasparov が、コンピュータは悪手を打たないことを信じて序盤戦を対局したことにある[29]。しかし、第 2 局の Deep Blue の手で Kasparov は判断を誤り、最終的には Deep Blue が勝利したのである。また自動運転車の開発も重要なステップである。この関連では、先に触れた米国の研究機関 DARPA が再びその役割を担っており、例えば 2004 年には DARPA グランドチャレンジが初めて開催された。この競技は自動運転車のみを対象とし、最速かつ自律的にレースを完走した車両が勝者となるものである。2005 年のネバダ州で開催された全長約 212km のレースでは、Volks Wagen 社製の”Stanley”が時速 35km で、6 時間 53 分 58 秒で優勝した[30]。2004 年の第一回大会では自律走行車がゴールに到着することすらかなわなかったことに鑑みれば、まさにこの結果はブレイクスルーといえるものであった[31]。さらに 2006 年、CMU の自動車である Boss は米軍飛行場で交通規則を遵守しつつ安全に走行することに成功し[32]、そして 2009 年には初期モデルの Google Seif Driving Car がカリフォルニアの高速道路を走行し、2015 年には累計 160 万 km を公道走行するようになった[33]。もう一つの重要な事例は、IBM が開発した Watson である。この Watson は、2011 年に米国のクイズ番組である「Jeopardy!」のチャンピオン Jennings と Rutter の 2 人に勝利した[34]。このクイズ番組の特徴として、問題に正解するだけでなく、異分野の知識を複合させる必要があった。Watson は人間の司会者の質問を正しく理解し、いわば皮肉的な言辞をフィルタリングして、正解を準備し、かつ、人間の対戦相手よりも速くブザーを押すことができたのである[35]。あるプログラムが人間の相手を理解し、それに反応することが可能になったことは、AI の歴史における一種のマイルストーンと見なされるべきだろう[36]。しかし、Watson では、問題を読み上げながら処理し、データベースから適切な回答を検索し、それを瞬時に出力するという、ソフトウェアの圧倒的なスピードがその成功の決め手となったことも言うまでもない[37]。

そしてここ数年続いている深層学習に裏打ちされた第 3 次 AI ブームは、まだ沈静化したというわけではない。2016 年、Google DeepMind は囲碁プログラムである「Alpha Go」を開発したが[38]、このソフトウェアは 2016 年 3 月の対局で、当時世界チャンピオンの囲碁棋士

---

29 伊庭斉志『ゲーム AI と深層学習　ニューロ進化と人間性』（オーム社、2018 年）5 頁以下。

30 *Russel/Norvig,supra* (fn.17), p.28.

31 ローレンス・D・バーンズ・クリストファー・シュルガン（児島修 訳）『AUTONOMY 自動運転の開発と未来』（辰巳出版、2020 年）49 頁。

32 *Russell/Norvig*, *supra* (fn.17), p.28.

33 その詳細は、日経 EXTEC「Google 社の『Waymo』が自動運転開発に与えるインパクト」（2016 年 12 月 26 日）を参照されたい。

34 松尾・前掲（注 21）19 頁参照。

35 金山博・武田浩一「Watson：クイズ番組に挑戦する質問応答システム」情報処理 52 巻 7 号 840 頁以下。

36 金山・武田・前掲（注 35）842 頁。

37 金山・武田・前掲（注 35）849 頁。

38 松原仁・伊藤毅志「AlphaGo の技術と対戦」人工知能 31 巻 3 号（2016 年）441 頁。

であった Lee Sedol に勝利した[39]。その理由は、過去の膨大な対局データのみならず、ソフトウェアが自らと対戦することで常にその実力を向上させることができたことにあるという[40]。

このように、AI の歴史を概観すると、AI は自己改善するシステムだけでなく、部分的に人間とのコミュニケーション能力を持つシステムや、人間に新たなソリューションを示すシステムも含まれていることがわかる。しかしその利活用の多様性は、「AI」自体の不明確さをももたらしている。AI の利活用における（刑）法上の問題を議論するにあたっても、検討の対象とすべき AI を明確にするために、まずはその「AI」の定義を明確にしなければならない。そこで次節ではこのことについて一定の指針を与える。

第 2 節　AI の定義への試み

AI の定義は多義的であり、例えば、「人間の思考プロセスと同じような形で動作するプログラム、あるいは人間が知的と感じる情報処理・技術」[41]（総務省）や、「インテリジェントな機械、とりわけコンピュータプログラムを作るための科学技術。それは人間の知能を理解するためにコンピュータの類似のタスクを用いることと関連付けられるが、生物学的に観察されうる方法でそれ自体を制限する必要はない」[42]（John McCarthy）などが挙げられる。このように、「AI」概念には統一的定義が存在するというわけではない。むしろ、この概念の統一的定義を確立しようとすること自体非常に困難ともいえよう。これは、ダートマス会議で McCarthy が紹介した“Artifical intelligence“という概念の解釈が不正確だったことが原因とされる[43]。そこで様々なアプローチが試みられた。

まず、AI を人間の知能と対極の存在として定義する試みがなされてきた。これによると、人間の知的なものとして見なされる能力それ自体を発揮するシステムを人工知能と呼ぶことになる[44]。しかし、人間の態度がどこから知的であるかを判断する統一された基準がないため、この定義はあまり意味を為さないようにも思える。とはいえ、概念定義が存在しないとなると、どのようなものを知的と見なすのかについて個別に異なる基準を設定することが重要となる。例えば、数学の素人にとっては、どんなに複雑な算数の問題でも頭の中ですぐに解けるような人のことを「知的である」というかもしれない。しかし、そのような計算能力だけで、あるシステムを AI システムとして理解するのであれば、それは否定されなければならない。なぜなら、この定義では単なる電卓でも AI という名称を与えなければならないこととなるが、それは間違いなく AI ではなかろう。

---

39 日本経済新聞「囲碁 AI、プロに 4 勝 1 敗　最終局も熱戦制す」（2016 年 3 月 15 日）

40 松原仁「コンピュータ囲碁の進歩」日本ロボット学会誌 35 巻 3 号（2017 年）192 頁。

41 総務省「令和元年度版 情報通信白書 第 1 部 第 3 節 2 .AI に関する動向 (1)」（2019 年）

42 *McCarthy*, WHAT IS ARTIFICIAL INTELLIGENCE?, Computer Science Department, Stanford University, 2003. p.2.

43 *Herberger,* „Künstliche Intelligenz“ und Recht, NJW 2018, S. 2825 (2826).

44 *Kaplan,* Artificial Intelligence: What Everyone Needs to Know, Oxford, 2016, p.7

これに対して、情報学者の Rich は、「現時点で人間が得意とすることをシステムが行うことができれば、人工知能と呼ぶことができる」という点に着目して定義を試みる[45]。高速な計算能力を持つシステムは、その点において我々人間よりはるかに優れており、逆にこの分野で優れているのは人間ではないため、一見すると、この定義は当を得たものであり、さらに、我々人間が多彩なシステムより優れている分野は未だ存在しうるので、この定義は当面の間は有効なものともいえる。もっとも、この定義を別の側面から見ると完全なものではない。例えば Alpha GO が、自分自身と対戦することで常に自らの実力を改善し、人間に勝利したが、この定義では、現在ではすでに人間より優れているという理由で、この囲碁システムは AI であるとされなければならない。しかし、自主学習によって達成したシステムなので、どのようにして思考過程を形成したのかも考えなければならず、この点が従来のシステムとは大きく異なる点である。そのため、Rich の定義が必ずしも AI システムを正確に分類できるわけではない。

さらなる AI の定義の試みとしては、数学者 Turing が提唱したアプローチであるチューリング・テストがある。システムがインテリジェントであると言うために、彼はその定義ではなく、合格しなければならないテストを試みたのである。彼が模倣ゲームと呼ぶこのテストでは、被験者はコンピュータを用い、相手の声を聞くことも見ることもなく、もう 1 人の被験者及び機械とコミュニケーションをとる。そこで被験者は、いくつかの問題を投げかけることによって、会話の相手のうち、どちらが人間でどちらが機械なのかを区別する。5 分間の会話時間の後、少なくとも 30%以上の被験者が、本来は機械である方を人間だと回答すればテストに合格したとみなされ、機械はインテリジェントであるということができる[46]。

このテストの支持者は、チューリング・テストを用いると、AI を複雑かつ不正確な定義に当てはめる必要がないため、極めて容易に AI を特定することが可能であるといった利点があるという[47]。というのも、これに合格するためには、AI は自然言語を処理することが可能であり、すでに述べられたものを再述するための膨大な知識量と記憶容量を持っているだけでなく、会話パートナーに適応することができなければならないからだとする[48]。それゆえ、チューリング・テストでは、自然言語の使用と処理、および未知の状況への適応能力が、AI と判定されるための決定的な基準となる[49]。しかしこのテストは、AI の特定領域にのみと関連するものにすぎないし、そもそもこのテストに参加するためには、まず AI がチャット機能で人間相手にコミュニケーションすることが可能でなければならない。そのような機能を持たない AI（例えば Alpha Go）は、テスト対象として最初から除外される。したがって、このテストは、AI の存在を確認するというよりもむしろ、いわゆるチャットボ

45 *Rich,* Artificial Intelligence, McGraw-Hill, New York, 1983, p. 3.
46 *Russell/Norvig, supra* (fn.17) p. 984.
47 *Lenzen,* Künstliche Intelligenz Was sie kann & was uns erwartet? , C.H. Beck 2018, S. 25.
48 *Russell/Norvig, supra* (fn.17) p. 2.
49 *Lohmann,* Strafrecht im Zeitalter von künstliche Intelligenz, Nomos 2021, S.44,

ット[50]の適性判断に相応するものであると言わざるをえない[51]。また、AI に付随しうる検索機能が、AI としての決定的要素になり得るかどうかも考慮しなければならない。入力された大規模なデータベースではパターン検索のみが行われる[52]が、この機能は一般に AI にとっての最低条件と言える。仮に大規模なデータベースがなければ、AI は学習に際して困難を要し、このようなテストがインテリジェントな機械として特定するのに適切なのかどうかが問題となる。人間の知性もまた、一つのテスト結果だけでは描ききれないほど多くの側面を持っており、このことは、機械知能のテストにも同様に当てはまる。しかし、チューリング・テストではあくまで自然言語処理と適応能力という 2 つの機能のみがテストされるため、画像処理などその他の側面は依然として考慮されない。したがって、1 つのテストですべての AI を特定できるわけではない。

このように、AI の定義の確立においては、それぞれのアプローチにメリットとデメリットがあるため、普遍的な定義は存在しない。このことは、AI という言葉がいかに政治、社会、メディアで日常的に使われているかを考えれば、その現状と相反するようなものに思われる。そのため、定義が困難な AI 概念をより明確なものにするために、次節では、AI の現象形態を見ることにする。

第 3 節　強い AI と弱い AI

現在、AI はさまざまな形で利用されている。例えば、産業分野だけでなく、医療機器[53]や刑事裁判[54]でさえも AI の存在が見受けられる。もっとも、AI が突然、産業技術から医療、確定判決に至るまであらゆるものをコントロールできるようになったというような、一見すると圧倒的な力を持つように見えるということには、より詳細な検討が必要である。このような個別の適用方法ごとに、どのような種類の AI が問題となるのかを区別することが重要である。そのような AI の現象形態としては、「弱い AI」と「強い AI」の 2 種類に分けることができる。この分類は人間の認知能力をコンピュータ上で再現するにあたり考慮すべ

---

50 チャットボットとは、人間と機械だけが対話に入るが、人間のコミュニケーションを可能にする対話システムまたは支援システムである。

51 *Calo,* People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship, Penn State Law Review 2010, S. 830.

52 *Matthias,* Automaten als Träger von Rechten, Logos 2008, S. 220.

53 例えば、いわゆる黒色皮膚がんを発見するための医療診断支援に AI を利用される可能性が指摘される。*Haenssle* u.a., Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists, Annals of Oncology, Vol. 29, Issue 8, 2018, pp.1837 参照。

54 米国では Eric L. Loomis に禁固 6 年の判決が下された。この判決は、「Compas」という AI ソフトに基づいたものである。最大の問題は、裁判所も Loomis も、この判断に至ったアルゴリズムを立証できなかったことである。しかし製造者は、その背後にあるアルゴリズムを公表しておらず、取引上の秘密であると説明している。*Liptak* (May 1st, 2017)*,* Sent to Prison by a Software Program's Secret Algorithms, New York Times を参照。

き AI の分類である[55]。

「弱い AI」とは、厳密かつ精確な方法で仮説を立て、検証することを可能にする道具にすぎない AI である。例えば、先述した Alpha Go のみならず、 Siri（音声認識）、 Google 画像検索（画像認識）など、特定分野に限って利用できるものである。

それに対して、「強い AI」とは、正しくプログラムされたコンピュータが認識状態を持ち、そのプログラムが人間の認識を説明する AI である。この AI に基づいて正しくプログラムされたコンピュータは知能となるとされる[56]。すなわち、心理的説明を検証する道具ではなく、むしろそのプログラム自体が説明となる。

この強い AI の開発は目下進行中のものである。これは、我が国の研究者の尽力に代表されるものであり、人間に近似したロボットの開発が進められている[57]。ここでは、外観の類似性[58]に価値を置くのみならず、ロボットの内面的な価値も重要視されている。そして、人間との機能的な一体化を目指し、感情を認識し、共感を示すものでなければならないとされる[59]。例えば石黒浩教授は、演劇の観客もしくは講義を受ける学生が、誰が人間で誰が機械なのかを確認するのに困難なほど人間に類似したアンドロイドの開発を行っている。彼によると、子供から始まり、最後は老人に至るまで、完全な自律型ロボットの開発までを行うとする。これらはいずれ、多様な状況で人間と自由にコミュニケーションをとる能力を有し、その結果、外見上だけでなく、その行動の結果として人間と同レベルのものと類似するようになるという[60]。

このような展開は、強い AI の開発がもはや遠くないことを示すものともいえるが、現在の技術状況を考えた場合「強い AI」は存在しないといわざるをえない。そのため、本稿の議論の対象とすべきは「弱い AI」である。

第 4 節　汎用型 AI と特化型 AI

汎用型 AI と特化型 AI の分類は、人間のように広範な課題を処理できるか否かに関する AI の分類であり、「強い AI」と「弱い AI」の分類とは一線を画する。

「汎用型 AI」とは、ある目標や状況からの知識を別の目標や状況に対して汎化する「転移学習」により目的や状況の変化に自ら適応する広範な能力をもつ AI である[61]。この点に

---

55 *Searle*, MINDS, BRAINS, AND PROGRAMMS, University of California, 1980, p.2.

56 *Ibrd*.

57 *Asada*, Towards Artificial Empathy. International Journal of Social Robotics, Vol.7, No.1, pp.19-33, 2015

58 例えば、浅田研究室が開発した人間そっくりの幼児ロボット”Afetto”は、その表面が赤ちゃんの肌のような感触だとされる。*Asada Research Group*「子供アンドロイドの開発」(http://www.er.ams.eng.osaka-u.ac./jpasadalab/?page_id=177)（最終アクセス 2022 年 11 月 28 日）

59 *Eberl,* Smarte Maschinen: Wie Kuenstliche Intelligenz unser Leben veraendert, HANSER, 2016, S. 291.

60 石黒共生ヒューマンロボットインタラクションプロジェクト（https://www.jst.go.jp/erato/ishiguro/outline.html）(最終アクセス 2022 年 11 月 28 日)参照。

61 ゲーツェル・ベン「汎用人工知能概観」人工知能 29 巻 3 号（2014 年）228 頁。なおこの用語は, *Gubrud*, “Nanotechnology and International Security”, Fifth Foresight Conference on Molecular

ついて「強い AI」に存在する「認識の有無」は問題としないところに注意しなければならない[62]。

他方、「特化型 AI」とは特定のタスクを行うように設計されているものであり、その限りにおいて人間以上の能力を発揮するが、それ以外のタスクは解くことができない AI である[63]。これに当てはまる AI の例としては前述した Alpha Go の他に、ELIZA, Deep Blue, Tay, Alice&Bob が挙げられる。まず、「ELIZA」とは Joseph Weizenbaum（MIT）による簡単なパターンマッチング技法を使った自然言語処理プログラムのことである。患者役のユーザーが入力するスクリプトに対して、心理療法士を装う「DOCTOR」が来談者と診断会話をシミュレーションするものであった。次に「Tay」とは、Microsoft 製のチャットボットで、Twitter 上にそれを公開し、不特定多数のユーザーとのコミュニケーションを通じて言語を習得するものであった。しかし、公開後 24 時間以内で差別的発言を繰り返すようになり、わずか 1 日で公開停止となってしまった。さらに、「Alice & Bob」とは、Facebook 製のチャットボットであり、実験段階で「両者」の会話は通常の用語法では理解できないやりとりを始めたため、公開停止となったことで知られている。

現在の技術水準に照らすと、存在が認められるのは専ら「特化型 AI」である。そのため、本稿の対象とすべきは「特化型 AI」である。

第 5 節　AI と学習

ここでは、AI 分野における「学習」概念について俯瞰する。より具体的には、機械学習と深層学習、そして特別な学習形態としてのフリート・ラーニングが挙げられる。

第 1 款 機械学習

機械学習とは、「明示的にプログラミングすることなく、コンピュータに学ぶ能力を与えようとする研究分野」であるとされ、さらにこれには「教師あり学習」、「教師なし学習」、そして「強化学習」という 3 つの学習手法が存在する。

まず、「教師あり学習」とは、事前に与えられたデータをいわば「例題（先生からの助言）」とみなして、それをガイドに学習を行う手法である。これは、入力されたデータに対して正しい出力を返すことを学習の目標にするものであり、徐々に入力と出力の関係を学習するものである[64]。例えば、犬の写真 10 枚を「これは犬である」として見せ、次に猫の写真 10 枚を「これは猫である」として見せる。その後にいずれかの写真を見せて、それが何かを当てる課題が挙げられる。

Nanotechnology(November 1997)が初出であるとされる。

62 鳥海不二夫「人工知能技術を俯瞰する」立法と調査 405 号（2018 年 10 月）5 頁。

63 鳥海・前掲（注 62）5 頁。

64 谷口忠大『イラストで学ぶ 人工知能概論 [改訂第 2 版]』（講談社、2020 年）205 頁。

次に「教師なし学習」とは、学習データに正解を与えない状態で学習させる手法である。その理由は、入力データがあらかじめ構造化されてシステムに提示されるのではなく、システム自身が構造化してパターンを認識しなければならないことにあり、クラスタリングは、このような学習の典型例である。これは、新たな知見を得るため、あるいはそもそもシステムが利用できるようになった（入力）データのパターンを認識するために利用される[65]。例えば、100 人の漫画のキャラクターを見せ、そのキャラクターの類似性に基づいて 10 グループに分ける課題が挙げられる。

さらに「強化学習」とは、学習データに正解はないが、目的として設定された「報酬」を最大化するための行動を学習する手法のことである。ここでは、適切な目標を設定した後に環境と相互作用し、その結果が正しければ、一種の報酬である強化を獲得することでその経験から直接学習するシステムである[66]。システムは最終的に行為の結果を記憶し、報酬を受け取ることを目的に行動しなければならない[67]。そのため、自分の行動経過を分析し、どの行為が報酬もしくは罰となるかをフィルタリングしなければならない[68]。これを繰り返すことによりシステムの機能は改善されていく[69]。例えば、迷路を抜けた時にのみゴールに到着したことがわかり、それを褒められるが、途中の経路は教えてもらえないような学習が挙げられる。

第 2 款　深層学習と人工ニューラルネットワーク

伝統的な学習形態とは対照的に、近年登場した AI の学習形態として、人工ニューラルネットワークを用いたいわゆるディープラーニング（深層学習）がある。これは、「概念の階層から、コンピュータは、単純な概念から複雑な概念を構築することにより、複雑な概念を学習することができる。これらの概念がどのように相互に構築されているかを示す図を描くと、その構図は深く、多くの層がある。このため、このアプローチを AI 深層学習と呼ぶ」[70]と説明される。

深層学習が機能するためには、いわゆる人工ニューラルネットワークが必要とされる[71]。人工ニューラルネットワークは、脊椎動物の脳のような実際の神経回路網を再現しようとするものである。モデルとする人間のニューロンは約 10^{10} 個あり、脳内での情報伝達に中心的な役割を果たす[72]。細胞体から軸索が伸び、その先端の終末ボタンという部位がその隣の細胞体から出る樹状突起とシナプス結合することにより、一つのニューロンから他のニ

---

65 松尾・前掲（注 21）118 頁。

66 *Russell/Norvig, supra* (fn.17), p 789.

67 *Sutton/Barto,* Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning series), Bradford Books, 1998, p. 3.

68 *Russell/Norvig, supra* (fn.17), p.789.

69 谷口・前掲（注 64）129 頁。

70 *Ian Goodfellow and Yoshua Bengio and Aaron Courville.* Deep Learning, The MIT Press, 2016, p.1.

71 小林・前掲（注 26）138 頁。

72 小林・前掲（注 26）138 頁。

ューロンへと信号を伝達する。このとき、各ニューロンはシナプスと結合された他のニューロンから微弱な電気信号を受け、その総和がある閾値を超えると興奮状態となり、信号はシナプスを通じて他のニューロンに伝達される[73]。このプロセスが人間の高度な情報処理製造者メカニズムを担う。この機能は、人工ニューラルネットワークに引き継がれ、そこでは、ニューラルネットワークを正確に再現しようとすることなく、抽象的にモデル化する試みがなされている[74]。例えば、階層型ネットワークはノード（ニューロン)、エッジ（シナプス)、および複数の層で構成され、各層は階層的に接続されており、各層はその上下の層としか接続されていない。入力層は情報を受け取り、その重みに応じてそれを隠れ層に伝達する。

深層学習を可能にするためには、システムはまず、その訓練段階でそのタスクを処理できるようにしなければならない。これらの学習手順は、さまざまな方法で実施することができる。教師あり学習では、システムに入力値を与え、システムが出力値を計算する。そして、これらの出力値が実際に正しい出力値からどれだけ誤差があるかを検査し、人工ニューラルネットワークの重みと接続を調整することで誤差を最小化する。この方法は誤差逆伝播法とも呼ばれており、主にパターン認識の分野で応用されている[75]。

深層学習方式のもう一つの種類は、いわゆる畳み込みニューラルネットワーク（Convolutional Neural Network）であり、これらは、畳み込み層、プーリング層、全結合層、出力層で構成されている[76]。ここでは、入力画像に対してカーネル（重みフィルタである線形行列）を畳み込み処理することで、そのカーネルに対応した特徴マップを得る。次に、特徴マップをプーリングすることで特徴マップのサイズを縮小する。これらを通じて、入力画像の微小な位置ズレや回転などを補正することができる。この畳み込み処理とプーリング処理を繰り返し適用し、特徴マップを抽出する。抽出した特徴マップを全結合層へ入力し，最終的に各クラスの確率を出力する[77]。このネットワークでは、画像認識の各タスクに合わせた出力層を設計することで、画像分類だけでなく物体検出も同時に行うことができる[78]。

深層学習の手法は、主にパターン認識の領域で用いられており、利用可能なデータ量が多いほど深層学習は有効に機能する。システムに膨大な量のデータが与えられるならば、深層学習方式は誤差に対応しうるという利点がある。つまり、人工ニューラルネットワークに何らかの異常が生じても、それがシステム全体の崩壊には至らず、結果としてシステム全体の性能が向上するため、個々の異常を補填することができる[79]。深層学習は、主に自動運転の領域で、周囲の環境認識のために用いられつつある。

---

[73] 小林・前掲（注 26）138 頁。

[74] 谷口・前掲（注 64）229 頁。

[75] 小林・前掲（注 26）143 頁。

[76] 谷口・前掲（注 64）233 頁。

[77] 藤吉弘亘「機械学習の進展による画像認識技術の変遷」計測と制御 58 巻 4 号（2019 年）293 頁

[78] 藤吉・前掲（注 77）294 頁。

[79] *Lenzen*, a.a.O. (fn.49), S. 63.

第 3 款　フリート・ラーニング

AI のもう一つの学習方法は、いわゆるフリート・ラーニングである。フリート・ラーニングとは自動運転領域に由来する概念であるが、この場合のエンドデバイスは特定の車両である[80]。AI が機能するためには、大量のデータが必要とされ、訓練段階中にもそのデータが利用可能な状態になければならない。しかし、このデータを利用するだけでは、過去のデータが再現できるにすぎないという問題を伴う[81]。そこで、AI による将来の決定が過去の学習内容に影響されないようにするため、現在の状況をその判断に反映させるよう、定期的にデータを更新することが試みられなければならないが、その可能性のひとつが、いわゆるフリート・ラーニングである[82]。その背景にあるのは、学習と改善を可能にするように、当該システムはクラウドからデータを取得し、そこに新たなデータを常にアップロードするものの、そのようなアップロードが低帯域などのために技術的に不可能な場合が多いことについて問題視されていることである[83]。そこで、フリート・ラーニングという手法により、一定のネットワークのもと、データをクラウドに転送する必要なく、エンドデバイスそれ自身で学習させることを可能にする[84]。具体的には、あるデバイスでローカルに学習したモデルを集めて 1 つのモデルに統合し、最終的に再びすべてのデバイスに同期して共有するので、車両団が他の車両のデータまたは経験から利益を得ることができる。この学習プロセスにより、個々のシステムの継続的な自己改善が達成されるだけでなく、集団内の車両がすでにデータを収集し、伝達されている範囲では、少なくともデータの最新性も保証される。これは、それぞれのシステムが得た知識を交換し、協働するという形で、既知の問題を解決することを目的としている。ここでは、全体の問題をアルゴリズムで分解し、個々のシステムが個々の小さな問題を解決し、発見されたソリューションによって最後に再び全体の問題を解決する。しかし、フリート・ラーニングの場合、各車両が自らの下位問題を解決し、これらソリューションが最終的に道路上で安全に動作することを達成するために一括して集計されるわけではなく、個々の車両が他の車両の学習記録から全体としての利益を得る[85]。

なお近年では同様の用語として、いわゆる Car-to-Car communication（車両間の通信）や Car-to-X communication（他のシステムとの通信）もある[86]。この通信形態では、ネットワーク接続された車両が、他のネットワーク接続された車両から情報を受信する[87]。この車間通

---

[80] *Fraunhofer IAIS*, Maschinelles Lernen „on the edge“, S.2.
[81] *Lenzen*, a.a.O. (fn.49), S. 61.
[82] *Lohmann*, a.a.O. (fn.59), S.51
[83] *Fraunhofer IAIS*, a.a.O. (fn. 80), S. 1.
[84] *Fraunhofer IAIS*, a.a.O. (fn. 80), S. 1.
[85] もっとも、収集したデータにおいては利用者の位置情報が含まれることもあり、場合によっては当該利用者の個人情報としてみなされることもありうる。この点、当該データの取り扱いにつき、分析を施すなどプロファイリングを行うことは欧州では規制される。しかし、この論点については本論稿の対象からは外れるため、その検討は別稿に譲りたい。
[86] *Jeschke,* Auf dem Weg zu einer „neuen KI“: Verteilte intelligente Systeme, in: Jeschke/Isenhardt/Hees/Henning (Hrsg.), Automation, Communication and Cybernetics in Science and Enigneering 2015/2016, p. 499
[87] *Mercedes Benz,* Car-to-X Communication. Mercedes-Benz is starting a Europe-wide cooperation project.

信は、WLAN と同様の無線技術で動作する[88]。この新種のネットワークをつうじて、道路交通はより強固で安全なものになるとされる[89]。これは、緊急車両の接近、渋滞の解消、逆走運転に関する情報、路面凍結などの道路状況など、事故データから経路に関連する情報を、運転手自身がこれらを目視する前に受信し、反応できるようにするためにある。さらにこの技術は、いわゆる Road Side Unit、すなわち、信号や道路工事、現在の交通渋滞に関する情報を提供するインフラ内のデジタル機器との通信に使用することができる[90]。この新しい形式のコミュニケーションによって、道路交通が改善され、交通安全が向上し、運転がより快適になることが期待される[91]。

このような車両間のコミュニケーションは、AI の一領域なのか、それとも単なる情報伝達の一形態なのかが問題となるが、正しいのは後者である。車両間の通信や車両と環境との通信は、AI の一領域には該当しない。その理由は、個々のシステムは常に新しい情報を受け取り、それに従って作動するものの、例えばオイルが漏れた道路に近づくと、運転手がすぐに対応できるように車両が警告を発するなど、システムは他の車両や環境との相互作用を通じて学習するように訓練されていないからである[92]。従って、Car to Car Communication という概念は、一見するとフリート・ラーニングと同義に見えるとしても、それと同一視してはならない。AI の一領域と言えるのは、フリート・ラーニングの分野だけである。

第 4 款　二種類の AI システム

先述した学習プロセスにおいて、現在 AI システムには、クローズ AI とオープン AI 二種類があることを考慮に入れなければならない[93]。前者は、訓練段階が終了した後は学習を継続しない AI システムのことを指すので、これらのシステムは訓練段階を終えた後に到達した水準にとどまっている。その一方で、オープン・システムの場合は、これらも製造者側から訓練を受けるが、訓練段階が終了した後も学習を続けるという点で状況が異なる。このことは、製造責任者が訓練段階から製品をリリースし、各利用者と具体的に学習し続けることを意味する。このようにして、個別の利用者に最適化された AI 製品となるが、その一例が、Apple 社が開発したアシスタントシステムの「Siri」である。「Siri」は、利用者から新たなスキルを学ぶわけではないものの、個々の操作や習慣によって、利用者個人に最適化される。

---

https://group.mercedes-benz.com/innovation/case/connectivity/europe-wide-cooperation-car-to-x.html(最終アクセス 2022/05/31).

88 *Car 2 Car Communication Consortium,* Deployment of V2X communication based on IST-G5. https://www.car-2-car.org/fileadmin/press/pdf/CAR_2_CAR_Communication_Consortium_Statement_ITSG5.pdf (最終アクセス 2022/05/31)

89 *Jeschke*, a.a.O. (fn.86), p.499.

90 *Car to Car Communication Consortium,* C2C-CC Manifesto, 2007, p.31.

91 *Car to Car Communication Consortium,* Clear benefits for road safety and traffic efficiency. (https://www.car-2-car.org/about-c-its/, 最終アクセス 2022/05/31)

92 *Lohmann,* a.a.O. (fn.49), S.51.

93 この分類は *Lohmann,* a.a.O. (fn.49), S.60 による。

もっとも自動運転分野では、オープン AI はまだ使われておらず、クローズ AI のみが用いられている。

第 5 款 人間と AI のインタラクション

近年、自動運転や医療などの領域では、さまざまな AI 技術の浸透が進んでいる。その理由は、そのような領域でこそ、AI が特に効果的に人間をサポートし、高い成果を上げることが可能であるとされ、人間のリスク要因は AI によって最小化されるという。

しかし、それが効果的な方法であるように見えても、あらゆるケースで AI が人間の影響から切り離されて動作するわけではないので、たとえば欧州 AI ハイレベル専門家グループや日本の総務省は、人間が AI を監督するシステムを提案している[94]。この提案の背景には AI ハイレベル専門家グループが「信頼できる AI(Trustworthy AI)」の確立を望んでいることがある。監視やインタラクションが可能となるために検査できるからこそ、人間が信頼する AI であり、換言すると、信頼できるのは、人間より先に行動する AI ではない。人間が監督するモデルでは、人間が AI を観察し、場合によっては介入する可能性を与えなければならない[95]。それは、ある意味で人間と AI との協力関係を創ることになる。

これを実現するためには、HITL モデル、HOTL モデル、そして HIC モデルという 3 つの可能性がある。まず、HITL（Human in the Loop）モデルには、このようなシステムがすでに稼働中であるときに人間があらゆる決定に介入できるような協力が含まれる。“in the loop“という表現にあるように、人間は AI システムの一部である。AI の利点を十分に活かせるかどうかは別個の問題ではあるものの、AI は個別の適用領域の改善に資するべきであることは言える。しかし、人間が常に介入すると、改善は遅々として進まないか、あるいは、およそ改善さえなされないこともありうる。

これに対して、いわゆる HOTL（Human On the Loop）モデルとは、人間がシステムの設計サイクルの形成に専ら従事しつつ、システムの作動中にのみ監督できるモデルである。ここで人間は、開発段階ではシステムが機能するように障害を取り除くが、設計段階が終了した後は一歩後退し、単なる監督者としての役割を果たすことになる。このモデルは、自動運転の分野でも頻繁に見受けられる。例えば、深層学習を行うカメラを搭載した車両で様々なテスト走行を行う際、この開発段階において「誤学習」を回避するために、カメラが学習したことを確認し、場合によってはそれを消去する形で人間が介入する。このようなシステムの開発プロセスを経てカメラが車両に搭載されると、人間は単なる観察者の役割を担うだけであり、システムは人間の影響を受けずに自立的に動作するようになる。

さらに、いわゆる HIC（Human in Command）モデルは、AI システムが計画領域で使用可

---

94 総務省「AI ネットワーク社会推進会議 報告書 2021～『安心・安全で信頼性のある AI の社会実装」の推進～』（2021 年）102 頁。

95 以下のモデルの説明は、*HLEG on AI* (High-Level Expert Group on Artificial Intelligence), Ethics Guidelines for trustworthy AI, 2019, p. 16 を参照。

能かどうか、使用可能な場合は個別状況に応じてどのように利用されるべきかを、人間が（予備的に）決定するものである。例えば、AI システムが融資の可否を決定するようなものが挙げられる。人間がまず、個別状況に応じて AI システムを利用するどうか、利用する場合はどのようにするかを決定する。つまり、人間がそのシステムを利用することを決定した限りで、AI システムは人間によって正確なパラメータを与えられ、これらに基づいてシステムが独自に判断する。そして、その決定を受容もしくは拒否するかを責任者である人間が選択することができる。

これらの説明は、AI は自立的に行動する可能性はあるものの、必ずしも人間の存在が無視されているわけではないことを意味している。

第 6 款　小括

上記で紹介した学習プロセスは、AI が様々な方法で学習可能であることを示している。それに応じて、最終的に AI が利用される分野によって、異なる学習プロセスが目的に適したものとなる。AI は、これまで人間が行っていたタスクを引き継ぎ、人間に比肩できないほどのスピードでこれを実行することで、多くの領域で人間の生活を便利にするとされる。このように AI は利便性を有する反面、ブラックボックス化というデメリットも無視することはできない。ブラックボックス化とは、AI の学習プロセスは分かっていても、個々の学習プロセスにおいて、なぜそのような結果になるのか、人間には理解できない場合が生じることである。特に人工ニューラルネットワークの領域では、これらのネットワークの内部で一体何が起こっているのか、とりわけ AI がノードになぜある重み付けをしたのかを確実に説明することはできない[96]。ニューラルネットワークの領域ではこの不明確性から、AI はブラックボックスとも呼ばれる[97]。まとめると、AI は、設定された課題を解決するためのアプローチが必ずしも人間には理解されえないという点で、学習に人間のコントロールが及ばないのである。

このような性質を抱えた AI の利活用をつうじて、本論文の議論の対象でもある、人間の生命・身体・財産が侵害された場合における刑事責任の検討に当たり、たとえ「弱い AI」ないしは「特化型 AI」であるとしても、その学習手法によっては人間には予期し得ないような思考過程を構築することがありうる。このことは、結果との因果関係における不明確性をもたらしうるため、次節では、まさにこのような問題意識のもと、AI の特性を考慮しなければならないとされる事例を概観する。

第 6 節　AI の利活用と刑法上の問題

AI が搭載された製品は、その利用者が詳細には予測できない判断を下すこともありうる。例えば、自動運転車は、予測できない、部分的なスマート環境と協働しなければならず、そ

---

96 *HLEG on AI*, *supra* (fn.95), p. 13
97 *Ibid*.

れに搭載される AI は、内蔵されている学習方式に応じ、そのダイナミックな環境に対する作業予測を確立するために、保存されているすべての情報と入力された情報を使用する。利用者は、AI が自身で取得した情報を独自に分析し、それに応じて自立的に作動することを知っているはずである。そのような AI の特性を認識した、その利活用に関わる各主体に対し、発生した法益侵害結果を、刑法上どのように負責させるべきか、それとも負責すべきでないのか。この問題が関連する事例は網羅的に検討する必要がある。それにあたり、まずは AI に起因する人間の生命・身体・財産の法益侵害事例について概観する。

第 1 款　自動運転車—アシャッフェンブルグ事例と東名高速事例

海外の運転支援システムの事例については、ドイツのアシャッフェンブルグ事例が挙げられる[98]。2012 年春、高性能の車線維持システムを搭載した自動車が、アシャッフェンブルク近郊のアルゼナウ村に進入したが、その入口で、運転手の 60 歳前後の男性が脳卒中となり意識を失った[99]。その際、彼はハンドルを右に切り損ね、通常であれば村に入る前に茂みの中で停止するところであったところ、車線維持システムが車を道路に戻すように誘導したため、車は高速でアルゼナウ村に突っ込み、女性とその子供が死亡した。その子供の父親はジャンプすることで助かり、足を負傷しただけであった。この事件は、自律システムが人間の生命を「侵害した」事例であると考えられよう。民法上は、ドイツ道路交通法 7 条の危険責任のために、特別な問題が提起されることはないが、より困難なのは刑法上の評価である。この事例においては、死亡事故を過失により惹起していない事故車両の運転手は答責的ではなく、2 人の人間の死と 1 人の身体傷害に基づく過失致傷もしくは過失致死の答責主体として製造者が考慮されるという。

日本では、運転補助システム（レベル 2）における交通事故（横浜地裁令和 2 年 3 月 30 日: 判例秘書 LLI/DB L07550489）[100]が発生している。後に検討を行うが、さしあたり簡単にその事例の概要を紹介する。これは、2018 年 4 月 28 日、神奈川県綾瀬市内の東名高速道路上において、被告人はレベル 2 の普通乗用自動車を、運転支援システムを使用して走行中、仮睡状態に陥り、前を走行中の車が車線変更した後もそのまま進行し、進路前方に停車していた普通自動二輪車に加速した状態で衝突し、それにより、同自動二輪車を前方に跳ね飛ばして前方に佇立していた被害者 3 名に衝突させて、そのうちの 1 名を死亡させ、2 名に傷害を負わせた事案であった。被告人には、運転中止義務違反に基づく過失運転致死傷罪により禁錮 3 年（執行猶予 5 年）が言い渡され確定した。

---

98 *Hilgendorf*, a.a.O. (fn.2), S.66.

99 この運転手が死亡したか否かは不見当である。Vgl. *Hilgendorf*, a.a.O.(Fn.2), S.68.

100 この判例を検討したものとして、中川由賀「具体的事故事例分析を通じた自動運転車の交通事故に関する刑事責任の研究②　～運転支援車(レベル 2) の事故～」中京法学 55 号 1 巻(2020 年)4 頁以下や、樋笠尭士「自動運転（レベル 2 及び 3 ）をめぐる刑事実務上の争点—レベル 2 東名事故を手がかりに—」捜査研究 847 号（2021 年）46 頁がある。

この事例はいわゆるレベル 2 の自動運転車[101]による交通事故における利用者の注意義務を明らかにした初の判例であり、2010 年代より盛んに議論されてきた自動運転車の事故事例をめぐる刑事責任の検討に大きな示唆を与えるものとなろう。もっとも、この検討から発展して、もし当該自動車がレベル 3 相当のものであったり、まだ実用化されてはいないがレベル 4 相当のものであったりする場合には事情は異なる。というのも、2020 年道路交通法改正および 2022 年道路交通法改正においてレベル 3 ないしはレベル 4 相当の自動運転車に関する定義規定、及びそれに伴う利用者（特定自動運行実施者）の義務規定が創設されたからである。むろん改正法の規定に従った、自動運転車の利用による利用者、ひいては製造者等に課せられる義務、及びそれに伴う刑事責任の帰属の妥当性に関しては別途検討を要する。

第 2 款 介護ロボット

近時の日本では、介護人材の不足が大きな課題となっている。介護分野の人材を確保する一方で、限られたマンパワーを有効に活用する解決策の一つとして、高齢者の自立支援を促進し、質の高い介護を実現するための AI を搭載したロボット・センサー等の活用が期待される。その中で、介護ロボット[102]を利用する老人ホームも見られるようになりつつあり、さらに、このようなロボットは将来的には、家庭のヘルパーとしても使われるようになるだろう[103]。

このような介護ロボットの利活用における刑法上の問題として、例えば、AI を搭載した介護用ロボットが多世代で暮らす家庭に販売されるという、架空ではあるが、前述のように将来的には十分ありうるケースを想定する。例えば、ある家庭では、物忘れがひどくなってきた祖母 A のために、このロボットが定期的に飲み物や食べ物を運搬しており、その間母親 B は別の仕事に専念していた。この AI ロボットが作動していたとき、予期しないセンサーの誤作動により毛布の上で遊んでいた B の子である C を轢いて死亡させてしまった。しかし、このセンサーは、製造企業による定期メンテナンスが行われており、これまでそのような不具合は知られておらず、予期もされていなかった。

この事例において製造企業や利用者（管理者）[104]の過失による刑事責任を検討するにあたっては少し注意が必要である。まず、製造企業に対する過失犯の成否につき、この AI ロボ

---

101 以下、自動運転車のレベルに関しては特に断りのない限り、SAE 基準を用いるものとする。

102 「介護ロボット」とは、ロボット技術が応用され利用者の自立支援や介護者の負担の軽減に役立つ介護機器である。その種類としては、移乗支援、移動支援、排泄支援そして認知症の方の見守りなどが挙げられる。（厚生労働省ホームページ https://www.mhlw.go.jp/file/06-Seisakujouhou-12300000-Roukenkyoku/0000210895.pdf 参照(最終アクセス 2022 年 11 月 28 日)）

103 日本における介護ロボットの現状の利活用例については、厚生労働省「介護ロボット導入活用事例集 2021」（2021 年）https://www.mhlw.go.jp/content/12300000/000928395.pdf (最終アクセス 2022 年 11 月 28 日) を参照。

104 上記事例では B を指すものとする。

ットの誤作動に対応する刑法上の注意義務の有無については、当該誤作動に対する予見可能性、そしてこの誤作動に対応する作為義務を事故当時に有していたか否かが問題となる。そして利用者（管理者）に対しては、この AI ロボットの誤作動による結果発生を予見し得たか否かが問題となる。

第 3 款　産業用ロボットーバウナタール事例

工場生産に供される産業用ロボットにも AI を搭載しているものはあり、それが人間の生命・身体を侵害することもありうる。実際に、次のような事例がドイツであった。

バウナタールの工場棟で、2015 年 6 月に作業員 A が遮蔽されていないロボットアームにつかまれて死亡した。他の作業員 B が、誤って規定よりも早くこの機械のスイッチを作動させてしまったという[105]。CNN でも報道されたこの事故のニュース[106]は、ロボットによる人間の殺害と関連するものにみえるかもしれない。もっとも、B が注意義務に違反したことが証明できれば、過失致死罪が成立し得る。ここ数年間でロボットが遮蔽された場所で働くのではなく、人間と直接接触することができるようになったため、この種の事例はありえないものではなくなってきている。さらに、部門長や企業の安全管理者などが、労働者に対して保障人的地位にあった場合には、そのような安全管理者の発生結果に対する負責も考えられる。しかしながら、この事件で大きな注目を集めたのは、人間が自働するロボット（機械）に殺害されたように見えたことによる。これは、人間に危害をもたらしうる「人工的な存在」への深層的な恐怖や空想に訴えかけるものであったとされる[107]。

この事例は、AI が搭載されていない産業用ロボットが人間の生命を侵害したものではあるが、仮にロボットアームが AI の自立判断により作動するものであったとするならば、本稿で検討する問題に沿うものとなる。すなわち、バウナタール事件の変形事例として、ロボットアームがもっぱら自立学習を行う AI の統制のもとで作動していたものとすれば、A の死亡結果と B の行為の間にある因果関係の機序が不明確となりうるからである。さらにこの場合、B の背後にいる管理者の保障義務の検討についてはなお慎重になされなければならないだろう。

第 4 款 過失犯処罰規定のない犯罪類型

さらに AI を搭載した機械がネットワーク上で、コンピュータ犯罪や、経済犯罪の構成要件が実現されることがありうる。特に、経済領域においては、近時ますますそのサービスが展開されているとされる AI・アルゴリズム投資における、相場操縦やこれを介したインサ

---

105 https://www.nha.de/kassel/kreis-kassel/baunatal-ort3125l6/toedlicher-roboter-unfall-bei-vw-kassel-in-baunatal-vor-gericht-8413531.html （最終アクセス 2022 年 11 月 28 日).

106 https://edition.cnn.com/2015/07 /02/europe/germany-volkswagen-robot-AI-kills-worker/index.html（最終アクセス 2021 年 9 月 30 日）.

107 *Hilgendorf,* Autonome Systeme, künstliche Intelligenz und Roboter Eine Orientierung aus strafrechtlicher Perspektive, FS Fischer, S.107.

イダー取引のような証券犯罪、さらにはデジタル・カルテルのような競争法違反が挙げられる。これらの犯罪において問題となるのは、利用者ないしは製造者の予期しない AI の学習結果によって当該犯罪の構成要件が実現された場合の刑事責任の所在である。これらの犯罪類型には、過失犯処罰規定が存在しない。そのため、学習しない AI・アルゴリズムの使用であれば利用者ないし製造者にその刑事責任が帰属しうるところ、学習を行う AI が介入することによって、彼らの故意の認定が困難になるのではないかということである[108]。

また、コンピュータ犯罪の領域では、学習によって利用者に最適化された AI に対する個人情報のハッキングをはじめとする不正アクセス罪（不正アクセス禁止法 3 条・8 条）の成否や、そのような AI の利用を妨害する場合の電子計算機損壊等業務妨害（刑法 234 条の 2）、不正指令電磁的記録作成・供用罪（刑法 168 条の 2）の成否、さらには学習を行う AI を搭載した電子商取引に供されるソフトウェア・エージェントに対して不正アクセスがなされ、虚偽ないしは不実の電磁的記録により当該ソフトウェア・エージェントの管理者の財産が騙取されたが、その電磁的記録が不正アクセス時における虚偽の情報もしくは不正な指令の供与によって作出されたのか、それとも不正アクセス後の AI の学習結果によって作出されたのかが不明確な場合における電子計算機等使用詐欺罪（刑法 246 条の 2）の成否[109]などが考えられる。このように、人間が行為主体で、AI が行為客体となる類型を中心にした議論があまり見られないところ、これらの事例はサイバーセキュリティ上の観点から不可欠なものであるといえる。

第 5 款　小括

以上のような AI 製品に起因する人間の生命・身体・財産等の法益侵害についての刑法上の責任帰属問題は実際の事案から仮想事例まで含めて喫緊の課題ともいえよう。まず、従来の自動運転車の刑法上の議論について、レベル 2 相当の自動運転車については実務上の見解が示されており、同時に二度の道路交通法改正を経てレベル 3 ないしは 4 の自動運転車に関する利用者の義務が詳細に明文化されている[110]。もっとも、上記で示したようなアシャッフェンブルク事件[111]や東名高速事故のようなレベル 2 相当の自動運転車に関する利用者

---

108 このことについては、第 3 章で詳細に検討を行う。

109 この場合、不正アクセス罪は成立することを前提にする。

110 この状況はドイツでも同様である。例えば、StVO（ドイツ道路交通法）第 8 次改正において新設された 1a 条における「高度な又は完全な自動運転機能を有する自動車」とはレベル 3,4,5 の自動運転車を指し、1b 条以下でこれら自動車の利用者・製造者に対する義務を名文化している。その概要については、泉眞樹子「ドイツにおける自動運転車の公道通行 ―第 8 次道路交通法改正―」国立国会図書館（2018 年）https://dl.ndl.go.jp/view/download/digidepo_11052071_po_02750004.pdf?contentNo=1(2022/05/31 閲覧)を参照。

111 もっとも *Hilgendorf*, a.a.O, (Fn.2), S.68 では、利用者は「すでに被害惹起者としての明確に帰属可能である行為が欠如し、そもそも注意違反の存在を指摘することができない」という理由で刑事責任の検討から利用者を排除し、専ら製造者の刑事責任のみを検討している。しかし、なぜ利用者に帰属可能な行為が

や製造者の法律上の義務に関しては依然として普通自動車（レベル 0）のそれと同一であるが果たしてこの状況は妥当なものであろうか[112]。近時の動向を踏まえて再度検討を要するものと思われる。

その一方で、AI を利活用した製品は今や自動運転車に限定されず、介護現場で利用される介護ロボットや、製品工場においてオートメーションに供される産業用ロボットなど実体を持つ AI から、AI・アルゴリズム投資のような実体を持たない AI も存在し、そこにも刑法上の問題をはらむので、自動運転車をベースにした議論が隆盛を極めていた頃に比べて、この問題はより多彩なものとなり、より具体的なものとなってきた。これらの問題は断片的にではなく、体系的な検討をすることが必要であり、このことこそが、AI 製品の利活用における刑法上の問題を再び仔細に検討する意義である。次章では、そのなかでも、従前の議論の対象であった生命・身体侵害に対する検討を行う。その際、過剰な刑事責任を課すことにより AI の利活用や技術開発を萎縮させないように配慮しつつ、AI 製品の挙動の予見不可能性を前面に出すことにより、法益侵害結果が、いかなる人間的主体にも帰属されなくなるという「帰属の間隙」は可能な限り生じるべきでないという立場から検討を行う。

第 2 章　AI 製品の利活用における刑法上の諸問題―生命・身体への侵害事例

第 1 節　問題の所在

AI 製品を市場流通させる場合、あらゆる製品の市場流通の場合と同様に、さまざまなアクターが現れる。その範囲は、製造者から個々の販売者、そして利用者にまで及ぶ。そこでまず提起される問題は、製造者が市場に流通させた AI システムに欠陥があった場合、どのようにして製造責任者に刑法上の責任を問いうるのかということである。特に当該製品搭載の AI システムについては、リスクのある行動をさらに学習させ、その結果、損害が発生した場合に、製造者が非難されうるかどうかが問われる。この問題の解決にあたり様々な試みがなされたことは先述の通りである。しかし近年では、人間が当該 AI 製品を監視するモデルを維持する要請がある。そうだとすると、どの程度まで利用者側、製造者側が AI に対してどこまで監視・管理をしなければならないのかを明確にする必要がある。

第 2 節　将来的な技術水準の AI 製品における具体的検討

本節ではまず、第 1 章第 6 節で示した事例を基調に考察する。ここで、前述した介護ロボット事例ではそのセンサーが周囲の環境を学習して障害物の有無を探知するものであり、その学習結果による判断で A を轢いたものとし、産業用ロボット事例では当該センサーが自立学習により最適化されるものだったところ、その学習結果による判断で B を人間と判

欠如するといえるのかの説明はなされていない。

112 免許制度に関する指摘ではあるが、運転支援システムの監視制御にかかる訓練・技能はレベル 2 の自動運転車から必要であると古川・前掲（注 16）176 頁の脚注 16 は言う。

断できず圧死させてしまったものとする[113]。これら２つの事例における各主体の罪責はどのようなものになるのか。

一見すると、介護ロボットの利用者 X には A の死亡に対する（重）過失致死が、運用者（Y）に関しては A の死亡に対する過失致死罪の成否が、さらに、産業用ロボットの管理者である W には B の死亡に対する業務上過失致死罪の成否が問題となる。しかしながら X,Y,W の注意義務の確定において、「AI 製品の予測不可能な動作」をその予見可能性の対象に入れてしまうと、過失の範囲が過度に広がる恐れがあり、利用者・運用者の負担が増加することになる。

そこで、AI 製品の製造者(Z,V)を考慮すると両者とも業務上過失致死罪の成否が問題となる。しかし、いずれの事例も必要な注意を尽くしていたにもかかわらず、刑事責任という負担を課してしまうと、却って技術開発を萎縮させてしまう恐れがあり、AI 製品の利活用・普が法によって阻害されるという帰結を導きかねない。この点において、当該事故を「不幸な事故」として処理する解決方法も考慮されるが、この帰結では法感情に背く可能性があるだろう。これら事例では必要とされる注意基準そのものが問題であり、この内容を明確にする必要がある。そこで、現状水準の AI 製品―運転支援システム―における先例を参照しつつ、AI 製品一般の問題に妥当するような指標を提示する。

第３節　現状の技術水準の AI 製品における具体的検討

本節では、AI を搭載した自動運転車（ここでは、SAE 基準レベル２相当の運転補助システム搭載の自動運転車）による実際の事故事例（**横浜地判令和 2 年 3 月 30 日**）[114]における利用者の刑事責任についての検討を行い、どのような注意義務が利用者に課されるのかを明確にするものとする。

第 1 款　事案の概要

被告人 X は、平成 30 年 4 月 29 日午後 2 時 44 分頃、普通乗用自動車を運転し、東名高速道路上り 29. 7 キロポスト付近道路を厚木インターチェンジ方面から横浜町田インターチェンジ方面に向かい進行中、眠気を覚え、前方注視が困難な状態に陥り、前記状態のまま運転を継続したことにより、同日午後 2 時 48 分頃、同市同自動車道上り 29. 2 キロポスト付近片側 3 車線道路の第 3 車両通行帯を進行中に仮睡状態に陥り、そのまま約 130 メートル進行し、同日午後 2 時 49 分頃、同市同自動車道上り 29. 1 キロポスト先道路において、自車進行通行帯の進路前方に停車していた普通自動二輪車の存在に気付かず、加速した状態で、同車後部に自車前部を衝突させ、その衝撃により同自動二輪車を前方に跳ね飛ばして同車前方に佇立していた A（当時 44 歳）並びに座っていた B（当時 43 歳）及び C（当時 44 歳）に順次衝突させて同人らを路上に転倒させた上、自車右後輪で前記 A を礫過し、よって、

---

113 この学習過程が事後的に解明できなかったものとする。

114 判例秘書 LLI/DB L07550489.

同人に頭部挫滅損傷の傷害を負わせ、即時、同所において、同人を同傷害により死亡させ、前記 B に全治まで 12 週間の加療を要する見込みの左足立方骨骨折等の傷害を、前記 C に全治 9 日間を要する頭部挫創等の傷害をそれぞれ負わせたというものである。

第 2 款　裁判所の判断

裁判所の認定およびその判断については以下の通りである。

①被告人の予見可能性

被告人 X の予見可能性について、「本件運転支援システムは、 自車前方の物体を検知できずに、 静止した車両と衝突しないようブレーキをかけたり減速したりすることができなくなる場合があるなど、いかなる状況においても適切に動作することを保証されたものではない」として、「被告人が、本件運転支援システムには自車を前方の物体の手前で停止させて衝突を回避する機能もあると理解していた可能性は否定できない」ものの、「被告人は、事故を防止する責任は基本的に運転者にあるという説明は受けており・・・被告人自身の認識としても、 本件運転支援システムを作動させていたとしても、 前方を注視し、 何かあったときには運転を替わるという意識を持ち続けて運転しなければいけないと思って運転していたというのであるから、 少なくとも、 本件運転支援システムが道路状況に応じた適切な動作をしないことがあり得ることは理解していたと認められる」とした。

さらに裁判所は、「本件運転支援システムでは、対応し難い事態 (例えば、他の自動車が、高速度で走行している自車の直前に急に割り込んできたり、 前方から逆走してきたりした場合や、 自車のタイヤがパンクするなどの故障が起きた場合、 前車の積み荷が路上に落下した場合等様々な事態が考えられる) が一般に起こりうることは明らかであるから、 被告人は、 本件高速道路という比較的本件運転支援システムを作動させるのに適した場所においても、 前方を注視して自ら適切に被告人車を操作しなければ、 本件運転支援システムでは対応し難い事態に対応できず、 事故を回避できない場合があり得ることを当然理解していたはずである」ので、「本件運転支援システムが道路状況に応じた適切な動作をせず、 又は本件運転支援システムでは対応し難い事態が生じたにもかかわらず、 被告人が仮睡状態に陥るなどして前方を注視できず、 被告人車を適切に操作しないことによって事故が発生して人が死傷する危険があ」ることを理由にして被告人の予見可能性を肯定した。

②注意義務違反と結果との因果関係

次に、注意義務違反については、「被告人車が午後 2 時 44 分頃に走行していた地点から本件事故の現場までの間に、 被告人車を本件車道にはみ出ることなく停車させることができる非常駐車帯が複数ありそこで一時休息したり同乗者と 運転を交替したりすることも可能であったことを踏まえると、 被告人は、 前記のとおり前方注視が困難になるほど強い眠気を覚えた時点で、直ちに運転を中止すべき自動車運転上の注意義務があったにもかかわ

らず、これを怠って前方注視が困難な状態のまま運転を継続し、運転中止義務に違反したと認められる」とした。

さらに注意義務違反と結果との因果関係については、「本件運転支援システムの機能は前方の物体との衝突を回避するために設計されたものではないのであるから、客観的には、被告人が自ら被告人車を操作しなければ被告人車が本件バイクに衝突する危険があったことは明らかであ」り、「そして、被告人が、本件運転支援システムについて、自車を前方の物体の手前で停止させて衝突を回避するように設計されたものであると理解していた可能性は否定できないが、・・・本件運転支援システムが道路状況に応じた適切な動作をしないことがあり得ることは理解していた」こと、加えて、「本件事故直前の被告人車は、前車が方向指示器を点滅させ車線変更を開始したが、未だ同車の車体の一部が被告人車の走行する第 3 車両通行帯にある時点から・・・加速し続けている・・・被告人車の挙動は、本件バイクの手前で余裕をもって停止する前の走行としては不合理なものであり、被告人がこれまで経験してきた被告人車の挙動とは大きく異なることは明らかであるから、被告人が仮睡状態に陥っていなければ、本件運転支援システムが、被告人車を本件バイクの手前で停止させるように動作していない可能性を認識できたはずである」とし、「仮睡状態に陥ることなく前方を注視していれば、・・・被告人車が本件運転支援システムの動作によっては本件バイクの手前で停止できずに衝突する危険を予見し、急制動の措置を採ることは可能であり、被告人がそのような措置を採っていれば、被告人車が本件バイクに衝突することを回避することができたと認められる」ので、「本件事故は、被告人による前記運転中止義務違反に基づく危険が現実化したものと認められるから、両者の間の因果関係も認められる」と被告人 X の運転中止義務違反と被害者 A、B、C らの致死傷の結果との間に因果関係を認めた。

第 3 款　検討

日本では、2020 年道路交通法改正により、利用者・整備点検者に自動運行装置の作動状態記録装置による記録（道交法 64 条の 2 の 2）が、利用者に対して当該自動運行装置に係る使用条件[115]の遵守（道交法 71 条の 4 の 2）が創設された。ただし、自動運転車の利用者の義務は、普通自動車の運転者とほぼ同様である。もっとも、この「自動運転車」は SAE 基準のレベル 3 以上のものを指すので、利用者（運転者）に対して課せられる注意義務は、普通自動車のそれと比較しつつ結果予見義務・結果回避義務を検討することとなるものと思

115 （道路運送法 41 条 2 項）
プログラム（電子計算機（入出力装置を含む。この項及び第九十九条の三第一項第一号を除き、以下同じ。）に対する指令であつて、一の結果を得ることができるように組み合わされたものをいう。以下同じ。）により自動的に自動車を運行させるために必要な、自動車の運行時の状態及び周囲の状況を検知するためのセンサー並びに当該センサーから送信された情報を処理するための電子計算機及びプログラムを主たる構成要素とする装置であつて、当該装置ごとに国土交通大臣が付する条件で使用される場合において、自動車を運行する者の操縦に係る認知、予測、判断及び操作に係る能力の全部を代替する機能を有し、かつ、当該機能の作動状態の確認に必要な情報を記録するための装置を備えるものをいう。

われる。

第 1 項　自動運行装置（運転支援システム）を利用する道交法上ドライバーの義務と過失

現行道路交通法のもとでのレベル 2 の自動運転車の利用者の義務やその他禁止行為は、普通自動車の運転者のそれとほぼ同様である。その義務は、本件に関するものならば、道交法 70 条（安全運転義務）に表れている。道交法 70 条は、道交法における運転者に課せられる明確な義務（たとえば、71 条各号所定の運転者の遵守事項、72 条所定の交通事故の場合における緊急措置義務、報告義務など）のみではまかないきれないものがあるため、これを補う趣旨で本条のような総括的かつ抽象的な規定が設けられたとされる[116]。

もっとも、道交法 70 条は非常に抽象的な文言であるため、明確性を欠き拡大解釈されるおそれがあるため、厳格に解釈されなければならない。そのような趣旨から、この条文により可罰性とされるのは、道路、交通、当該車両等の具体的状況のもとで、一般的にみて事故に結びつく蓋然性の高い危険な速度、方法による運転行為に限られるべきである（いわき簡判昭和 43 年 6 月 3 日下刑集 10 巻 6 号 635 頁）。また、過失による安全運転義務違反として処断するためには、過失によって、「他人に危害を及ぼすような速度と方法で運転した」ことを認定する必要がある（最判昭和 46 年 10 月 14 日刑集 25 巻 7 号 817 頁）。むろん、過失により本条の規定に違反し、その結果人を死傷させた場合は、過失運転致死傷罪（自動車運転死傷行為等処罰法 5 条）が成立し、本条の違反行為がその過失の内容となるとされる[117]。

もっとも、本条の罰則（117 条の 2 第 6 号）の適用にあたっては、比較的重い刑罰を定めているところ、当該行為が本条の違反行為としてその構成要件を充足しているか否かを十分に検討すべきである（大阪高判昭和 38 年 10 月 3 日高刑集 16 巻 7 号 550 頁参照）。本件事例のように、「居眠り運転」による交通事故の場合は、確かに道交法 70 条の安全運転義務違反の罪が成立する余地はあるが、このことが過失運転致死傷罪における「過失」の内容となるため、道交法 70 条違反の罪は個別には成立しない。そのため、本件においては過失犯の検討スキームに従い、注意義務を予見可能性・結果回避可能性の観点から考察すべきである。

第 2 項　注意義務の確定―予見可能性判断

被告人特有の事情として、⑴本件運転支援システムには自車を前方の物体の手前で停止させて衝突を回避する機能もあると理解していた可能性は否定できない、⑵事故を防止する責任は基本的に運転者にあるという説明は受けていること、⑶本件事故以前の経験から，本件運転支援システムが適切に動作して被告人車が停止する場合には，被告人車が，前方の物体の手前で急制動にならない程度の余裕をもって停止するものと理解していた、という

---

116 道路交通執務研究会編著（野下文雄原著）『執務資料 道路交通法解説〔18 訂版〕』（東京法令出版、2020 年）766 頁。

117 道路交通執務研究会・前掲（注 116）768 頁。

ものがある。このうち、⑴・⑶は被告人の予見可能性を否定することに傾く事情とも思われる[118]。

しかし判決は、「本件運転支援システムに対する被告人の認識を前提としても，本件運転支援システムでは対応し難い事態が一般的に起こり得ることは明らかであるから，被告人は，本件高速道路という比較的本件運転支援システムを作動させるのに適した場所においても，前方を注視して自ら適切に被告人車を操作しなければ，本件運転支援システムでは対応し難い事態に対応できず，事故を回避できない場合があり得ることを当然理解していたはず」であり、「いかなる状況においても適切に動作することを保証されたものではなく，本件事故当時の被告人車のマニュアルにおいても，運転者は，本件運転支援システムが作動していたとしても，常に道路に注意を払い，いつでも必要に応じて対応できるようにすることが求められていた」ことを理由に、運転支援システムでも対応できない不測の事態が発生しうるという認識は、普通自動車の運転者と同じく、当該被告人にも存在するとした。そうすると、自動運転車の運転者にも当然、このような予見可能性が認められることになる。

予見可能性の判断にあたっては、行為者が当該事故における「因果関係の基本的な部分」を予見することが可能であったか否かが問題となる。例えば、北大電気メス事件（札幌高判昭和 51 年 3 月 18 日高刑集 29 巻 1 号 78 頁）では、「結果発生の予見とは、内容の特定しない一般的・抽象的な危惧感ないし不安感を抱く程度では足りず、特定の構成要件的結果及びその結果の発生に至る因果関係の基本的部分の予見を意味するものと解すべきである。そして、この予見可能性の有無は、当該行為者の置かれた具体的状況に、これと同様の地位・状況に置かれた通常人をあてはめてみて判断すべきものである」とする。

しかし、本件事例のように、当該事故の原因の一つとされる、センサーの誤反応やそれに伴うブレーキシステムの誤作動に対するメカニズムを被告人は認識しているわけではない。このことについて、過失運転致死傷罪における死傷結果に至った当該自動車の誤作動のメカニズムは、被告人にとってはブラックボックス化していると言える。この「因果関係のブラックスボックス化」については、以下の事例が参考となる。すなわち、近鉄生駒トンネル火災事故上告審決定（最決平成 12 年 12 月 20 日刑集 54 巻 9 号 1095 頁）によると、「右事実関係の下においては、被告人は、右のような炭化導電路が形成されるという経過を具体的に予見することはできなかったとしても、右誘起電流が大地に流されずに本来流れるべきでない部分に長期間にわたり流れ続けることによって火災の発生に至る可能性があることを予見することはできたものというべきである」という。このように、仮に被告人が当該自動運転システムの誤作動のメカニズムを知らなかったとしても、予見可能性が肯定される余地がある。

さらに、因果経過のプロセスにおいては、本件事例のように、被告人の落ち度（この場合は居眠り）、センサー機器の誤作動、前車の挙動等種々の因果的推移が存在し、「因果経過の基本的な部分」が予見し難いこともありうるが、渋谷温泉施設爆発事故上告審決定（最決平

[118] 樋笠・前掲（注 100）51 頁。

成 28 年 5 月 25 日刑集 70 巻 5 号 117 頁）によると、「結果発生に至る因果のプロセスにおいて，複数の事態の発生が連鎖的に積み重なっているケースでは，過失行為と結果発生だけを捉えると，その因果の流れが希有な事例のように見え具体的な予見が可能であったかどうかが疑問視される場合でも，中間で発生した事態をある程度抽象的に捉えたときにそれぞれの連鎖が予見し得るものであれば，全体として予見可能性があるといえる場合がある」という。そのため、以上の判例のスキームに従えば、レベル 2 の自動運転車が予期せぬ挙動をした場合においても、その中間で発生する一般的な不測の事態そのものは予見可能であるとされ、予見可能性は認定されるものと思われる。

しかし、本件事例における予見可能性の具体的対象は、因果的経過を考慮すればむしろ、「事故を回避するために、自動車の異常を察知して、急制動措置を施すこと」にあるものと思われる。そうすると、被告人特有の事情である（3）を仔細に検討すべきであり、この点において、「普通自動車を走行するに当たり、一般的に仮睡状態に陥れば、交通事故を惹起しうること」とは枠組が異なるのである。

第 3 項　注意義務の確定—結果回避可能性

判決では、「被告人車が午後 2 時 44 分頃に走行していた地点から本件事故の現場までの間に，被告人車を本線車道にはみ出ることなく停車させることができる非常駐車帯が複数あり，そこで一時休息したり同乗者と運転を交替したりすることも可能であったこと」および「被告人は，仮睡状態に陥ることなく前方を注視していれば，本件バイクから約 19 メートルの地点までに，被告人車が本件運転支援システムの動作によっては本件バイクの手前で停止できずに衝突する危険を予見し，急制動の措置を採ることは可能であり，被告人がそのような措置を採っていれば，被告人車が本件バイクに衝突することを回避することができた」という。

過失犯においては、結果は行為自体の持つ許されない危険が現実化したものでなければならない。この過失犯の「客観的注意義務違反」を求めるには、行為者が物理的・生理的にまだ結果の実現を回避できる時点に求められなければならない[119]。例えば、最判平成 15 年 1 月 24 日（裁時 1332 号 4 頁：第二黄色信号点滅事故）によると、「本件は，被告人車の左後側部にA車の前部が突っ込む形で衝突した事故であり，本件事故の発生については，A車の特異な走行状況に留意する必要がある。…Aは，酒気を帯び，指定最高速度である時速 30 キロメートルを大幅に超える時速約 70 キロメートルで，足元に落とした携帯電話を拾うため前方を注視せずに走行し，対面信号機が赤色灯火の点滅を表示しているにもかかわらず，そのまま交差点に進入してきたことが認められるのである。このようなA車の走行状況にかんがみると，被告人において，本件事故を回避することが可能であったか否かについては，慎重な検討が必要である。」・・・「対面信号機が黄色灯火の点滅を表示している際，交差道路から，一時停止も徐行もせず，時速約 70 キロメートルという高速で進入してくる車両が

[119] 松宮孝明『先端刑法総論』（日本評論社、2020 年）135 頁以下。

あり得るとは，通常想定し難いものというべきである。しかも，当時は夜間であったから，たとえ相手方車両を視認したとしても，その速度を一瞬のうちに把握するのは困難であったと考えられる。こうした諸点にかんがみると，被告人車がA車を視認可能な地点に達したとしても，被告人において，現実にA車の存在を確認した上，衝突の危険を察知するまでには，若干の時間を要すると考えられるのであって，急制動の措置を講ずるのが遅れる可能性があることは，否定し難い。そうすると，…被告人が時速１０ないし１５キロメートルに減速して交差点内に進入していたとしても，上記の急制動の措置を講ずるまでの時間を考えると，被告人車が衝突地点の手前で停止することができ，衝突を回避することができたものと断定することは，困難であるといわざるを得ない。そして，他に特段の証拠がない本件においては，被告人車が本件交差点手前で時速１０ないし１５キロメートルに減速して交差道路の安全を確認していれば，A車との衝突を回避することが可能であったという事実については，合理的な疑いを容れる余地があるというべきである」という。このように、結果回避判断に際しては、特異な事情が介在している場合には、例えば、事故当時の道路状況や被害者ならびに被害車両の状況、さらにドライバー車両の特性等を詳細に検討すべきである。

裁判所は、「被告人車は，…本件バイクから約 45. 6 メートルの地点を時速約 13. 1 キロメートル（なお，この速度は，被告人車の走行状況や本件運転支援システムの作動状況等に関するログデータに基づき認定している。他の時点における被告人車の速度も同様である。）で走行していたが，前車との車間距離が開くことに対応して加速していった。さらに，被告人車は，午後 2 時 49 分 19 秒頃には，本件バイクから約 27. 5 メートルの地点を時速約 27. 8 キロメートルで走行しており，その時点では，前車はほとんど車線変更をし終わっていたところ，被告人車の前方には本件バイクを含め複数の車両が停車していたにもかかわらず，本件運転支援システムがこれらの車両を検知しない状況になったため，被告人車は更に加速し，午後 2 時 49 分 22 秒頃，本件バイクに衝突した…。そして，加速を開始してから本件バイクに衝突するまでの被告人車の最高速度は時速約 38. 1 キロメートルであるところ，この速度を前提として停止距離（空走距離と制動距離の合計）を計算しても，被告人が，仮睡状態に陥ることなく前方を注視し，遅くとも被告人車が本件バイクの約 19 メートル手前の地点に到達するまでに本件バイクとの衝突の危険を感知して急制動の措置を講じていれば，被告人車は本件バイクの手前で停止することができた」とする。

この点において、被告人側は「本件運転支援システムが搭載された自動車の運転手は，本件運転支援システムの機能を信頼しているため，自車が進路内の物体と衝突する危険が生じたとしても，本件運転支援システムを搭載していない自動車を運転する場合に危険を感知する時点で衝突の危険を感知するのではなく，本件運転支援システムに異常を感知した時点で衝突の危険を感知するのであり，本件事故においては，被告人車が，一般路走行時の通常の減速加速度では本件バイクの手前で止まれない距離まで減速を行わずに進んだ時点で，本件運転支援システムに異常を感知する」と言うが、裁判所はこの見解を採用していな

い。しかし、本件事故にける自動車はレベル 2 の自動運転車であり、普通自動車（レベル 0）の操作とは同一ではない。そのため、この点も含めて結果回避可能性の判断を為すものであり、この点を完全に捨象して考慮すべきではなかったと考えられる。

第 4 項　小括

本件事故におけるポイントは、注意義務の検討スキームにある。裁判所の示す注意義務違反は「一般的に仮睡状態に陥った場合に起こりうる事故を未然に防止する」という、運転中止義務がその内容であるが、本件事案ではむしろ「運転支援システムの異常を察知した時に急制動を施す義務」という義務であると考えられる。もっとも、いわゆる運転支援システムの誤作動に起因する交通事故において、仮睡状態に陥るといった、行為者にとって運転中止義務が認められる事例の場合は、この仮睡状態をもって過失を認定しうることは判示の通りである。しかし、このような事故事例においては、行為者の運転支援システムに対する「過信」が招来した事故であるとも言える。そうすると、政府の推進する自動運転実現の観点からしても、運転支援システムないしは自動運行装置を利用するドライバーには改めて、このような「過信」を起こさせないように啓発する必要性があろう。

ところで、本件事案における量刑事由として、運転支援システムの誤動作に関する事情は一切考慮されていない。仮に、当該行為者が仮睡状態に陥ることなくドライバーが運転した折に事故が発生した事例や、緊急時にのみドライバーに運転を引き継ぐレベル 3 の自動運転車で同種の事故が発生した場合でも、普通自動車におけるドライバーの注意義務をそのまま転用するのは妥当とは思えない。過失犯の処罰範囲をみだりに拡大しないためにも、注意義務（特に結果回避可能性）の内容や因果関係の認定にはより慎重な姿勢が必要であると思われる。

第 4 節　AI 製品の利活用による生命・身体侵害における刑法上の一般的考察

第 3 節で確認したように、AI 製品の利活用を通じた事故における刑事責任の確定をするにあたっては、ヒューマンエラーのみならず AI 特有の技術的観点も予見可能性や結果回避可能性ないしは因果関係の確定に必要であることがわかる。そうすると、AI が関与しない製品に起因する事故と同様の手法で検討するのでは不十分である。とりわけ AI 技術に関する部分では複雑かつ高度な専門知識が求められることもあるが、AI 製品を利用する主体や事件（事故）に関与する実務家がこれを把握するには困難を伴うことが多い。そこで、AI を利用する主体に対する明確な法的義務付けによって法的判断がなされるべきと考える。本節では、そのような義務付けに関し、現行法のもとで存在するものと存在しないものについて区別して、その各々について望ましい解決手法を提示することとする。

第 1 款　法的義務が存在するケース

法的義務は存在するものの例としては、SAE レベル 4 以上に該当する自動運転車をめぐ

る利用主体に関する法的義務が道路交通法（以下、「道交法」とする）に追加されたことは記憶に新しい。

この改正では、「特定自動運行」の定義（道交法 2 条 17 の 2 号）、特定自動運行の許可（道交法 75 条の 12）、特定自動運行計画の遵守義務（道交法 75 条の 18）、特定自動運行を行う前の措置（道交法 75 条の 19）、特定自動運行実施者の遵守義務（道交法 75 条の 20）、特定自動運行主任者の義務（道交法 75 条の 21）、特定自動運行終了時における義務（道交法 75 条の 22）、特定自動運行中の交通事故における措置（道交法 75 条の 23）[120]とレベル 4 以上の自動運転車の運行に際して、「特定自動運行主任者」や「現場措置業務実施者」という当該自動運転車を監視する主体を定義して、その運行に関する手順が事細かに規定されている。このように規定されるレベル 4 の自動運転車の運行にあたっては、人間による監視のもとでその利用を許可するスキームとなっている。ただし、個々で挙げられている主体による運行は「運転」概念には該当しないことに注意すべきである（道交法 2 条 17 号参照）。すなわち、レベル 4 以上の自動運転車の交通事故事例の場合は「特定自動運行主任者」や「現場措置業務実施者」が行為主体として考慮される。しかし、道交法 75 条の 21 には罰則規定が設けられておらず、加えて「運転」概念に該当しない以上、過失運転致死傷罪の適用はできない。そうすると業務上過失致死傷罪もしくは過失致死傷罪の適用の問題となる。これら過失の内容における注意義務は新設された特定自動運行主任者の義務などを手掛かりに構成することができよう。

これに対して、レベル 3 の自動運転車ついて検討すると、まず当該自動車の運転手の運転行為は道交法上の「運転」概念に該当する。そうすると、自動運行装置の利用者の義務としては、運転者の義務（道交法 70 条以下）の他に以下の義務が課せられる。すなわち、作動状態記録装置による記録等（道交法 63 条の 2 の 2）、自動運行装置を備えている自動車の運転者の遵守事項（道交法 71 条の 4 の 2）である。それでは、レベル 3 の自動運転車の交通事故の場合、刑法上はどのように処理されるか。先述の通り、運転者には道交法上の各義務違反および過失運転致死傷罪の適用が考慮される。このレベルの運転者に課せられる義務として、普通自動車の運転者に課せられるそれに加えて、オーバーライドに対応できるようにする義務がさらに（間接的に）課せられているといえる（71 条の 4 の 2 第 2 項）。しかし、道交法 71 条 5 の 5 号は、携帯電話用装置、自動車電話用装置その他の無線通話装置を通話のために使用し、又は当該自動車等に取り付けられ若しくは持ち込まれた画像表示用装置（後写鏡、窓拭き器その他の視野を確保する装置、速度計、走行距離計その他の計器、原動

[120] 罰則として、①1 項前段及び 3 項前段違反は、117 条 3 項：（人の死傷があった場合）5 年以下の懲役又は 100 万円以下の罰金、117 条の 5 第 2 項：（人の死傷がない場合）1 年以下の懲役又は 10 万円以下の罰金、123 条：両罰規定、②1 項後段及び 3 項後段違反は、119 条 2 項 6 号：3 月以下の懲役又は 5 万円以下の罰金、123 条：両罰規定、③2 項違反：117 条の 5 第 2 項：1 年以下の懲役又は 10 万円以下の罰金、123 条：両罰規定、④4 項違反は、120 条 2 項 4 号：5 万円以下の罰金、123 条：両罰規定である。

機付自転車の速度計除く）に表示された画像を注視しないことに限定されている[121]。なので、通話もしくは画面注視に限定されているため、例えば読書のように「画面以外のものを注視」する行為は対象外となり前方不注視として 70 条違反となるにすぎない。もっとも、道交法 71 条の 4 の 2 第 2 項や 71 条 5 の 5 号は普通自動車の運転よりも高度の注意をもって運転する義務を利用者に課していると考えられるが、特定自動運行実施者に関する義務と比較すると直接的には規定されていない。そう考えると、オーバーライドに対応する義務を規定した上で、除外要件を設定すべきではないだろうか。

第２款　法的義務が存在しないケース

第 1 款では法的義務が明文化されているケースについての考察を行った。その内容にはまだ検討すべき点はあるものの、AI 製品をめぐる主体の定義およびこれら主体の義務、そして考えられうる刑法上の問題とその解決に係る法律上の解釈の指針は明らかになったといえる。具体的には、レベル 3 ないしレベル 4 の自動運転車という AI 製品が交通事故により人を死傷させた事例が考えられる。レベル 3 の自動運転車ならば、道交法上はその利用者が道交法上の「運転者」と認められるので、過失運転致死傷罪の適用の問題となる。またレベル 4 の自動運転車ならば、利用主体として規定されている特定自動運行主任者や現場措置業務実施者が道交法上の「運転者」に該当しないため、彼らに対しては業務上過失致死傷罪ないしは過失致死傷罪の適用の問題となる。このとき、過失の内容たる注意義務違反における注意義務の根拠として新設された上記主体の義務を手掛かりに構成することができよう。しかし、刑事責任を導く根拠の一つとなりうる法律上の義務が存在しない場合、AI 製品をめぐる製造者や利用者にはどのように刑事責任は負責されることになるのか。この点については現行法の解釈論をベースに検討することにして、必要に応じて立法論を展開したい。

第 1 項　AI への刑事責任？

自動運転車における法整備がなされる以前は、自動運転車の事故事例を想定して、当該結果に対する刑事責任の帰属主体が欠如する可能性—帰属の間隙—や分散する可能性が指摘されてきた[122]。帰属主体が欠如するとなれば、自動運転車による事故が発生したとしても何人も刑事責任を負わないという帰結となり、自動運転車が普通自動車に比して特別視されることとなるが、その理由の説明に窮するどころか、法感情がそれを許さないだろう。また

---

121 この規定の罰則は 117 条の 4 第 1 号の 2 によると 1 年以下の懲役または 30 万円以下の罰金である。

122 主なものとして、*Beck*, Intelligente Agenten und Strafrecht. Fahrlässigkeit, Verantwortungsverteilung, elektronische Personalität, Studien zum deutschen und türkischen Strafrecht - Delikte gegen Persönlichkeitsrechte im türkischen-deutschen Rechtsvergleich (Band 4), Ankara 2015, S. 179 ff.（その紹介として、根津・前掲（注 8）105 頁）その他にも、根津・前掲（注 8）475 頁以下、稲谷・前掲（注 8）法律時報 91 巻 4 号(2019 年)54 頁、稲谷・前掲（注 8）日本ロボット学会誌 38 号 1 巻（2020 年）など。

帰属主体が分散するとなれば、自動運転車をめぐる主体（利用者、販売者、開発製造者）に対して常に刑事責任が帰属される可能性にさらされながら自動運転車に関わることとなる。これら状況下においては自動運転車の開発・普及を阻害する可能性がある。そこで帰属主体を一つに集約させる試みとして「AI」に対する刑事責任の帰属が展開されてきた[123]。

第１目　先行研究の素描

日本における議論の素描としては以下のようなものが挙げられる[124]。例えば、「伝統的な刑法学の立場からは、刑罰とは自然人のように、一個の人格が想定される存在にのみ科すことができるものとされている。このような観点からは AI に刑事責任を問う前提として、そもそも AI に人格を認めることが可能であるのかが問題」[125]とし、「AI が真の意味で我々の社会の対等なメンバーであるとの認識が共有されない限り、AI に独自の刑事責任を問うという方向性は否定されるべき」という消極的見解がある。他方、「① 刑罰の目的として、応報刑論を偏重するのではなく、抑止刑論と社会復帰論の意義を再確認すること、②刑法の存在意義を、社会構成員の法益侵害の予防に求めること、③社会の構成員は、人間に限られる必要はない。人間の仲間（fellow）として人間にとって重要な法益を侵害しうるが故に、侵害防止が義務付けられ、その義務が履行できる能力を有する存在であれば、当該能力の源泉を自由意思（free will）と位置づけ、刑法の対象である社会構成員として評価されること、④ AI の機能ないし能力が更に発展し、自律的学習能力が高まり、フレーム問題にも一定の対処ができるようになると、AI も社会の構成員となること、⑤AI に対する刑罰は、法益侵害に寄与したアルゴリズムの改変等であり、AI の再利用（AI による法益侵害の再発防止と、その社会復帰）に資するものであるべきこと、⑥ AI に対する刑事裁判は、インターネット上で行われること」[126]と「AI に対する刑罰」の条件を提示しつつこれを積極的に肯定する見解がある。前者の見解においては、人間との同等性を見いだせるか否かという Seelmann の見解と親和的である[127]。さらに後者の見解とは議論の方向性は異なるものの、法人処罰の文脈に引きつけて、AI を法人として擬制し、法人自体の処罰を肯定しうる法秩序が存在する以上は法人たる AI に刑罰を科すという着想のもと、積極的に「AI に対する刑罰」を肯定する見解も存在する[128]。

また、「AI に対する処罰」を肯定する可能性は否定できないという立場の見解もある。これによると、「たとえば事故を起こした自動運転車両に『刑罰』を科す際、どこまでが一人

---

123 以下の分類は、川口浩一「ロボット・AI に対する刑罰をめぐる最近の議論」法律論叢 94 巻 4・5 号（2022 年）100 頁以下を参考にした。

124 *Lohmann,* a.a.O. (fn.49) , S.105 も同旨。

125 深町・前掲（注 8）209 頁以下。

126 今井・前掲（注 8）31 頁。

127 *Seelmann,* Zurechnung zu Künstlicher Intelligenz?, in: FS Reinhard Merkel zum 70. Geburtstag Teilband I, 2020, S.695-706.

128 *Quarck*, Zur Strafbarkeit von e-Personen, ZIS 04/2020, S183.

の人格なのかという問題が生じるものの、その範囲を有益性の観点から画定する余地もありうる。その一人の人格に対し、現行法は有効な刑種を予定していないが、再犯防止の観点からは再プログラミング措置が考えうるところであり、その措置を施す際にロボットや AI が苦痛を感じているように『見える』のであれば、この問題も克服しうる。再プログラミングという措置の技術的特性と、未だ犯罪を行っていないロボットや AI ないし、自動運転車両にも、刑罰内容と同様のアップデートを施す必要があるため、個人責任の原則との抵触が考えられるが、ロボット法という新たな制度構築や原則の修正によって、あるいは人格の範囲を改めて検討することによって、この問題もまた克服される余地はある」と[129]。しかし、「ロボットや AI に『刑罰』を科すことは、理論的には全く可能性が無いわけではないが、その『刑罰』は真に『犯罪に対する責任非難』という枠内で用いられているかはなお慎重に検討されるべきである。単なる被害への不安から、スケープゴートとしてロボットや AI に『刑罰』を科すのであれば、それは『刑罰』の意義を変容させてしまう」と留保されている。このように、「AI に対する刑罰」の意義・機能の面ではそもそも人間に対する「刑法」と大きく異なり、その導入を目論むとしても解決しなければならない多くの問題を抱えることがうかがえる。

私見としては、この「AI に対する刑罰」にシフトすると、本来であれば結果帰属されるべき行為者が「AI」を隠れ蓑にして帰属から逃れる余地を与える可能性が拭えず、結果として望ましい解決方法ではないように思われる。そもそも AI という被造物たる道具に対してそのような価値を付与する意義も不明確である。さらにいえば、仮に将来的に AI と人間が同等の存在として認められるという状況が到来したとしても未解決の問題は存在する。それは、AI の刑事責任主体性に関する検討と、AI に対する「刑罰」に関する検討が仔細に行われているのに対して、AI が刑法上の行為を為しうるのかという問題である。これを扱う論稿はほとんど見られない。そこで、AI が刑事責任主体性を満たす可能性は留保しつつも、Lohmann の議論に沿って AI の行為性の検討を試みる[130]。

第２目　AI の刑法上の行為可能性

構成要件の検討に立ち入るためには、まず、構成要件的評価の対象となる行為とは何かという行為概念に関する検討が必要である。たとえば無意識の状態での動作など、刑法上重要な行為でなければ刑法上の検討は必要ない。そうすると、AI が行為概念の前提を満たすことができるのかという疑問が提起される。以下では、行為論に関する議論を踏まえつつ、この問題につき検討する。

---

[129] 根津・前掲（注 8）法学新報（中央大学）125 巻 11 号（2019 年）495 頁。

[130] 米法ベースの同様の議論として、*Lima*, Could At Agents Be Held Criminally Liable? Artificial Intelligence and The Challenges For Criminal Law, South Carolina Law Review 69, 677, 2018, pp.681.

（1） 行為論概説

刑法上の行為には限界要素、基本要素、結合要素の 3 つの機能があるという[131]。限界要素とは、例えば睡眠中の寝返りもしくは反射動作のように、およそ犯罪の対象とならないものを行為から除く機能を指す。次に基本要素とは、刑法並びに刑罰法規で規定される、刑法上重要な全ての人間的態度を例外なく包含しうる論理的可能性を備える共通の基盤であることをいう。そして結合要素とは、構成要件に該当し、違法で、有責な、という後続する規範的評価を結びつけ、そしてこれら判断を先取りしないようにするものをいう。もっともこれら 3 つの機能をすべて具備するような行為の定義はほぼ不可能であり、伝統的に議論されてきた行為論のもとではいずれかの機能を犠牲にしなければならないものであった。しかし、その展開を完全に無視しえないため、行為論をめぐる議論を改めて概観することにする。

行為論の出発点は自然的行為論である。この学説の支持者は、行為を「意思に基づく身体の動静」[132]もしくは「外界に対する有意的態度」で「有意的態度による外界の変更（ある結果）の惹起あるいは不阻止である」[133]と定義する。そこでは、外界の変化が人間の身体的運動に起因し、それが運動神経を介し、これを動かしている心理的なものに達した場合を行為と定義する[134]。この行為論においては特に「有意性」が求められており、それゆえ反射的動作、睡眠中の動作、無意識下の動作ないしは強制状態における身体的運動を行為から排除することができる。そのため、行為の限界要素機能を満たす。しかし、例えば転轍手が遮断器を降ろし忘れていた事例のような忘却犯（認識なき過失不作為犯）の場合は有意性が存在しないと批判された。

これに対し、いわゆる目的的行為論は、「人間の行為は目的的活動の遂行である」という命題を基に、行為者が一定の目的を予定し、その達成に必要な手段を選択したうえで、目的の達成のためにその行動を計画的に統制することができることにその目的性を基づかせるものとする[135]。これは、故意正犯の背後の過失行為を、過失による不可罰の共犯ではなく、過失正犯そのものであるという判例の出現に伴い、これを説明するために、故意作為と過失・不作為はそれぞれ行為の構造が異なるということを前提にする[136]。より詳しく言うと、過失・不作為は、構成要件上重要でない結果に向けられる目的的行為であると説明する。しかし、不作為に対し、不作為者は結果に対して因果的ではないため、その因果経過を操縦することができず、結果として目的的にも行為できないと批判され[137]、過失に対しては、「法的に無意味な目的性を行為概念にもちこむことは問題である」[138]とされたり、過失行為は落

---

131 その淵源は、*Maihofer*, Der Handlungs Begriff im Verbrechen System, 1953, S.10 にある。

132 山口厚『刑法総論[第 2 版]』（有斐閣、2007 年）42 頁。

133 *von Liszt,* Lehrbuch des Deutschen Strafrechts, 1919, S. 116

134 中山研一『刑法総論』（成文堂、1982 年）138 頁。

135 Vgl. *Welzel,* Das Deutsche Strafrecht 11 Aufl., 1969, S.33.

136 松宮孝明『刑法総論講義[補訂第 5 版]』（成文堂、2020 年）51 頁。

137 福田平『全訂刑法総論』（有斐閣、2011 年）56 頁など。

138 団藤重光『刑法綱要総論（第 3 版)』（弘文堂、1990 年）98 頁。

ち度のある人間の行為自体を指したりするのであり、それは積極的な目的志向的行為ではない[139]と批判されてきた。そしてこの行為論に対して決定的なのは、第二次世界大戦後の西ドイツ基本法 103 条 2 項における罪刑法定原則の明文化に伴い、不真正不作為犯の説明に窮したことである[140]。

そこで、不作為の行為性の説明を試みるために社会的行為概念が主張された。この行為概念は、行為を「社会的外界に向けられた有意的態度」[141]や、「予測可能な、社会的に重要な『結果』の有意的惹起」[142]、「客観的に予測可能な社会的結果へと方向づけられた客観的に支配可能な態度」[143]であるなどと定義する。この行為概念にとっての長所は、少なくとも、作為と不作為、故意と過失など、刑法上重要な行為態様の形態を全て把握するという機能を果たしている点にあるとされる。このような行為態様は、社会的外界を変化させる[144]。しかし、例えば睡眠中の動作は、社会的外界を変更することもあるため、刑法上重要でないものであっても、実際には顧慮されるべき経過もなお社会的行為論では「行為」に含まれる。また、社会的行為論に対しては、社会的に耐え難いもので、それゆえに社会的にも重要なものである場合は処罰されるため、結局は、立法者が社会的に重要なものを決定するのではないかと批判される。

さらに、人格的行為概念とは、行為を「人格の表出」[145]や、「人格の客観化」[146]、「行為者人格の主体的現実化」[147]と理解する。特に Roxin の定義では、その態度が（人間の）「心的―精神的な行為の中心」に分類される場合に行為が存在するという[148]。注意すべきなのは、反射的行為などは、意思的に支配できず、それゆえに人格性の表出が存在しないので行為からは排除される[149]。もっとも、例えば日本刑法 39 条 1 項のように、心神喪失者にも「行為」が可能であると規定されているように、仮に統合失調症により心神喪失状態である者に対して行為者人格の主体的現実化（ないしは表出）が認められるかは疑問とする批判がある[150]。

なお、著名な裁判例として大阪地判昭和 37 年 7 月 24 日下刑集 4 巻 7･8 号 696 頁がある。裁判所は、「行為者のある外部的挙動がその者の行為と許価され得るのは、その挙動が行為

---

139 大塚仁『刑法概説 総論（第 4 版）』（有斐閣、2008 年）102 頁。

140 松宮・前掲（注 136）51 頁。

141 *Von Liszt/E.Schmidt*, Lehrbuch des Deutschen Strafrechts, 26 Aufl. 1 Band, S.153.

142 *Engisch,* Vom Weltbild des Juristen, 2. Aufl. 1965, S.38.

143 *Maihofer*, Der soziale Handlungsbegriff, FS-Eb. Schmidt, 1961, S.151.

144 佐伯千仞『刑法講義 総論（4 訂版）』（有斐閣, 1984 年）145 頁参照。

145 *Roxin/Greco,* Strafrecht Allgemeiner Teil, Band I, 6. Aufl. 2020, § 8, Rn. 31..

146 *Arthur Kaufmann,* Die ontlogische Struktur der Handlung, Skizze einer personalen Handlungslehre, FS-H. Mayer, 1966, S.79.

147 団藤・前掲（注 138）105 頁。なお大塚・前掲（注 139）104 頁以下では、より詳細な分析とともに、この定義から発展して、行為を「行為者人格の主体的な発現としての有意性に基づく身体的動静であり、一般人の認識判断によって、その社会的意味が認められるもの」と定義する。

148 *Roxin/Greco,* a.a.O (fn.145), § 8, Rn. 44

149 *Roxin/Greco,* a.a.O (fn.145), § 8, Rn. 44

150 松宮・前掲（注 136）63 頁。

者の意思によつて支配せられているからであつて、右の意思支配が存しない場合には行為も存しないと言うべきであり、ある行為が刑罰法規の構成要件に該当するか否かは、右法規によつて要求される規範に従つて行為者が自らの行動を統制し得る意思の働らき即ち規範意識の活動に基ずいてなされた行為を対象としてなされるべきであ」るとして被告人行為の行為性を判断する。もっとも、第二審である大阪高判昭和 39 年 9 月 29 日（判例集未搭載）では行為性を認めつつ責任能力が欠けるとしたことに留意しなければならないが、少なくとも実務的見解では、刑法上の行為といえるためには、行為者の「意思」や「任意性」といった意的要素が要求されることが窺え、自然的行為論と親和的である。

（2） AI の刑法上の行為可能性

行為概念を適切に定義するために決定的なものは、冒頭に挙げた 3 つの要素すべてを把握しなければならないということである。これまでのところ、これら要素を一つに完全に統合することができる学説はないものの、それぞれの行為概念の当否は差し当たって留保しつつ、それぞれの行為概念に従って AI が「行為」できるかどうかが決定的である[151]。

自然的行為論は、睡眠中の動作がそうであるように、行為をするための意思が全く形成できない場合にのみ、行為性が排除される。AI の場合は、それがある行為をする意思を形成できるかどうかが問題となる。例えば、人間が腕を伸ばすのは、その瞬間にそれを欲する決定をしたからであるという例で明確となる。もし、AI が腕を伸ばせるとすれば、それは AI 自身が決定したのではなく、アルゴリズムが、その瞬間に必要だと判断したからであろう。AI は外部的に決定された方法、すなわち利用可能なデータを基に行為するが、それ自体は行為する意思を形成するものではない。この点を自立学習に関連づけると、AI 自身がデータを収集し、それに基づいて行動可能となるという可能性を意味するにすぎず、AI がプログラミングから完全に切り離されるということではない[152]。

目的的行為論によれば、行為者が一定の目的を予定し、その達成に必要な手段を選択したうえで、目的の達成のためにその行動を計画的に統制することができることが求められる。行為が刑法上重要となるためには、目的に方向づけられた行為、すなわち目的に方向づけられた意思[153]がその行動のトリガーとなる。しかし AI は、そもそも行動を行うように意思を自立的に形成することはできないため、なお意思による、ある目的に方向づけられた行動をとることができない。AI は、事前に実装されたデータや新たに学習したデータをもとに実行するアルゴリズムに基づいて動作する。さらに現状の AI は、自立的に目的基準を設定す

---

151 *Lohmann,* a.a.O. (fn.49), S.121.

152 今井猛嘉「自動車の自動運転と刑事実体法―その序論的考察」『西田典之先生献呈論文集』（2017 年,有斐閣）525 頁は、AI に搭載されるシステムに学習機能が具備され、当初想定されえなかった行動選択がシステム上可能となると、AI の意思に基づく、それ自身としての振る舞いが決定されているとし、AI の行為性を肯定できるものとする。しかし、この説明においては当初の学習機能が人間によるプログラミングに基づいていることを看過している。

153 *Gless/Weigend*, Intelligente Agenten und das Strafrecht, ZStW 2014, S. 561 (572).

る段階には程遠く、それに応じて行動できるようにするための目的基準を必要とする[154]。目的が確立された場合にのみ、システムは対応するデータによって行動できるのである。

社会的行為論によっても、AI システム側の刑法上重要な行為を構成することには一定の疑義がある。ここで決定的なのは、社会的に重要な態度が存在することである。この学説において刑法上重要な行為とは、何人も社会的に重要な結果を有意的に惹起した場合に存在するが、AI による振舞いが社会的に重要な意義を有するか否かが重要な問題となる。将来的にこれが肯定されるのであれば、AI の行為も肯定できるかもしれないが、これが社会的に重要な意義を有するか否かについての一致は現在のところみられるかは明確ではなく、根拠に乏しいものとなる[155]。

人格的行為概念については、その主観面において問題が生じる。結局、「意的」要素が入り込むために、AI がその態度を支配するための意思を形成するか否かの問題に帰着し、結果として因果的行為論における当てはめと同様に AI の行為可能性は現状では否定せざるをえない。

実務上も、前掲・大阪地判昭和 37 年 7 月 24 日によれば、少なくとも行為者の「意思」や「任意性」が行為には要求されることが窺えるから自然的行為論の当てはめと同様に AI には意思・任意性を認めるには困難を伴う。

以上の検討により、AI の行為可能性を認めるのは非常に困難であるといえる[156]。

第 2 項　不規制による解決

第 1 章第 6 節で示し、本章第 2 節に修正した想定事例において、一見すると AI を搭載した機械が人間の意思から離れて動作しており、もはやその AI 製品の製造者や利用者は刑事責任を負いえないと考えられるかもしれない。それは、当該製品は、利用者にとって予見不可能な動作をしたため、過失犯における予見可能性を欠くものと考えられ、さらに、製造者側についても必要な注意義務を履行していたとするならば、もはや当該結果に対していかなる主体にも法益侵害が帰属されえないように見えるからである。もっとも前述の検討から、AI が刑法上の行為を為すというには非常に高いハードルが存在するが、むしろこのような AI の予測不可能性こそが注意義務を導くため、AI の挙動が一般的に予測不可能であることはなお、AI の製造者や利用者を免責しえない。これは動物園の経営者が虎を檻から放し、その虎が通行人を殺害したとしても、虎は動物であるから制御することができなかった

---

154 これについては、*HLEG on AI,*a.a.O (fn.95)の定義を参照。

155 See *Lima*, *supra* (fn.130), p.685.

156 付言すると、客観的帰属論でも同様のことがいえる。この場合、「行為」は構成要件に規定された結果が行為者に対して彼／彼女の仕業として帰属できることによって初めて定義できる。この「行為」とは、その者の物理的・身体的能力によって結果発生が回避し得たか否か、そして生じた結果及びそこに至る過程とその出来事に責任を有する者の態度の社会的意味によって決まる（松宮・前掲（注 136）53 頁以下参照)。このとき、AI の「態度」の社会的意味を論ずる際、社会的行為論におけるロジックと同様に、現状ではこれを有するか否かについては開かれたままである。

と主張することはできないのと同様である[157]。

AI に刑事責任を負責させることは、行為概念の観点から非常に困難であるため、一般的に AI の利用者を免責することは、事実上、例えば自動運転車の誤操作による被害者の傷害に対して誰も刑事責任を問われないことを意味する[158]。そのため、自らの行為が他人の生命、身体、財産などの刑法上保護される利益を侵害する可能性を予見できる者は、その行為を慎む責務を負う。したがって、先の動物園の管理者は、虎が解放された場合に人間に危害を加えることを予見した場合、虎を檻から解放することを控えなければならない。同様に、このことは潜在的に危険な製品にも当てはまる。例えば、自動車の製造者が、適切な注意を払って、その自動車のブレーキが悪天候時に作用することが信頼できないことを知りえたにもかかわらず、その自動車を流通に置いた場合、製造者は注意義務に違反することになろう[159]。これらの検討により、不規制による解決は妥当なものではないことがわかる。

第 3 項　厳格責任による解決

AI の背後者をねらいとする刑事責任の負責の可否につき、AI の製造者に危険責任（厳格責任）を導入することが有望な解決の糸口であるとされることがある[160]。例えば米国では、民事製造物責任における設計上の欠陥、製造上の欠陥、指示・警告の欠陥の立証がそれに該当し、より具体的には、製造者側は危険な欠陥を最小限に抑えるように製品を製造すること（製造上の義務)、製品から実質的な危険性を排除するために合理的な注意を払うこと（設計上の義務)、残存する隠れた実質的な危険性について消費者に警告すること（指示・警告上の義務）についての義務を有するものとされる[161]。厳格責任の文脈では、不可避的な危険な製品の製造者は、その製品に製造上の欠陥がないことを保証し、消費者に隠れた危険性を警告しなければならない[162]。米国では 30 以上の州が、警告義務や更新義務など、製品の販売後に生じる義務の一部を採用する[163]。しかし、販売後の義務は米国の各州間で統一されておらず、裁判所は製造業者に販売後の更新義務を課すことを繰り返し拒否してきたことに注意しなければならない[164]。

刑法上の厳格責任として考えられるものについて、ドイツでは、義務違反ではない許され

---

[157] *Gleß, Silverman, Weigend*, If Robots Cause Harm, Who Is To Blame? Self-Driving Cars And Criminal Liability, New Criminal Law Review, Vol.19, Number 3, 2016, p.420.

[158] アメリカ法の民事責任の問題については,例えば *John W. Terwilleger*, Navigating the Road Ahead: Florida Autonomous Vehicle Statute and Its Effect on Liability, 89FLA. B.J. 26, 27 (205).

[159] この検討については第 4 項で詳細に行う。

[160] 日本の議論については、例えば、経済産業省・前掲（注 13）64 頁参照。米国の議論については *Wagner*, Robot Liability June 19, 2018), S. 13 f., https://ssrn.com/abstract;3198764 (最終アクセス 2022 年 11 月 27 日) を参照。

[161] *Marks*, US Product Liability Law, 2 INT'L Bus. LAWYER 69 1998.

[162] *Owen*, Inherent Product Hazards, 93 Ky. L.J. 2004, p.377, 379.

[163] *Stilwell*, Warning: You May Possess Continuing Duties After the Sale of Your Product! , 26 REV. LITIG. 2007, p. 1035, 1037.

[164] *Smith*, Proximity-Driven Liability, 102 GEO. L.J. 1777 (2014), p. 1802-08.

たリスク創出につき、民法上の負責に加えて刑法上の負責も問題となるとされる[165]。そうしたドイツ刑法におけるリスク負責の例としては、完全酩酊の構成要件（ドイツ刑法第 323 条 a）や喧嘩闘争への関与（ドイツ刑法第 231 条）が考えられる。これは、行為者に過失責任がなくても「許されざる事象に身を置いた」ためにすべての結果について負責されるという *versari in re illicita* の帰属原理による[166]。この可罰性は許されない、もしくは、社会的に不相当で望ましくない（完全酩酊事例の）行為者によるリスク創出と結び付く[167]。これは日本法における、許されないリスク負責についての同時傷害の特例（刑法 207 条）の議論と類似する。完全酩酊の文脈と同様で、その喧嘩闘争に関与することがリスクのある態度として評価されることにより当罰性を帯びるのである。もっとも、危険責任的思想に基づく刑罰は例外的なものであるため、これを一般原理として導入するべきでないし、条文の根拠もない。

第 1 目　リスク負責を客観的処罰要件に位置付ける構想

上述したリスク負責について、上記のような状況があって初めて結果発生は客観的処罰条件として構成することが正当であるとする見解もある[168]。しかし危険な AI システムの市場流通について、結果発生を客観的処罰条件として捉えようとすると[169]、これは市場流通の不許容性に根拠づけられなければならないという問題に直面する。この場合、「十分な安全性を確保せずに危険な製品を市場流通に置く」という抽象的危険犯の創設を前提とすればよい。ここで、義務違反ではない AI システムの製造、流通そして利用は、許される社会的に望ましい活動に該当する。しかし、許された態度はすでに概念上、法的な態度予期に違背しえないため、規範妥当を刑法によって確証する必要はない。したがって、許された態度についての刑法上のリスク負責は、刑法の非体系的な枠組破壊として否定されなければならず[170]、当該問題の解決に厳格責任のスキームを用いることはなお適切ではない。

第 2 目　DPA による解決

厳格責任の関連では、「確率的な危険のシステムによる統制を重視し、情報提供と精神・開発体制の改善、被害者への補償などを自主的に行うことを検察官と約束し、その見返りに刑事訴追を免れるという制度」[171]の導入が近時注目されている。これは、DPA（Deferred

---

165 *Fetah-Moghadam*, Innovationsverantwortung im Strafrecht: Zwischen strict liability, Fahrlässigkeit und erlaubtem Risiko – Zugleich ein Beitrag zur Digitalisierung des Strafrechts, ZStW 2018, S.881.

166 その概念については *Roxin/Greco*, a.a.O. (fn.145), § 10 Rn.122.も参照。

167 その支配的構想への批判については *Roxin/Greco*, a.a.O. (fn.145), §23 Rn.7 ff.を参照。

168 犯人が酩酊状態において遂行した犯行が、責任原理に対する違反を妨げるために,過失として扱われるという可罰性を要求する反対解釈については *Roxin/Greco*, a.a.O. (fn.145), §23 Rn.9 ff を参照。

169 *Hilgendorf,* a.a.O. (fn.107), S.111.

170 Vgl. *Fetah-Moghadam*, a.a.O. (Fn.165), S.882.

171稲谷・前掲（注 8）日本ロボット学会誌 38 巻 1 号（2020 年）40 頁。その詳細として、稲谷龍彦「Society 5.0 における新しいガバナンスシステムとサンクションの役割（上）」法律時報 94 巻 3 号（2021 年）105 頁。

Prosecution Agreement）という、企業に対して原因究明に必要なあらゆる情報の提供や、必要に応じて再発防止に向けた具体的な改善などを義務付ける代わりに、関係者の訴追を一定期間延期する訴追延期合意を技術開発の発展に伴う社会の変容に伴い、刑罰体系・責任制度の枠組も変化すべきという思想[172]に基づいて導入しようとする試みである。その具体的内容として、①事故に関するステークホルダーの厳格刑事責任を規定し、事故が起きた場合に、②リスクの発現であれば情報提供と損害賠償、③不確実性の発現の場合には情報提供と事故調査委員会による調査への協力及び製品・サービスの改善や組織の改善を約束させて訴追を延期する、というものとして、②～③について不協力・不履行の場合には訴追して相当額の制裁金を科すと共に、認証取消のような厳格な行政制裁を併せて行うという。

しかし特に①に関して、厳格責任を基調とする制裁は従来の刑罰観とは異なるのではないか、DPA に倣って企業にのみ責任を課すということは、関係する自然人に対する刑事責任は排除されるのかという疑義が指摘されたが[173]、それに対して、「厳格な法人処罰や DPA がない状態で、これだけ複雑な科学技術社会の情報の非対称性にどのように対応するか疑問。また、責任主義はシステマティックなリスクについてはほとんど意味をなさない。複雑なシステムが原因で死亡事故が起きた際に、一見非難出来そうな誰かを処罰して問題解決したことにすることが、我々の科学技術社会を正常に前進させるのか疑問であり、問われ始めているのは、正に刑事司法制度の意義そのものであ」り、「例えば、ある企業システムに置かれた人が、その置かれた文脈に即して合理的な行為をした結果、『犯罪』を実行したとして処罰されたとしても、システム自体が温存されれば、その立場におかれた別の人によって違法行為は繰り返されうる。さらに、製品の品質に起因する深刻な事故に関して、責任者が刑事罰を科されたにもかかわらず、その後も品質不正が生じた会社の例などは、企業の構造を変えないと、人々の行為が変化しないことを象徴しているように思われる」という[174]。

すなわち、事故が起きたからといって、個人責任を追及すれば良いわけではないので、関連する自然人の法的責任については、その追及が有効な場合に限ってなされるべきで、厳格な結果回避義務を課してゼロリスクを強調し、個人責任に傾斜する現在の刑事制裁制度は、技術開発に関する強烈な負のインセンティブだとして、新たな責任制度の枠組を強調する。しかし、この構想には新たな制度（法人制裁制度）を導入する必要性が拭えず、それどころか責任や刑罰に対する観点の変化を要するものである。結果としてこの解決方法にはかなりのハードルがあるため[175]、再度製造者や利用者の過失責任を検討すべきであろう。

---

[172] *Hilgendorf,* Können Roboter schuldhaft handeln?, in: *Hilgendorf/Beck*, a.a.O (fn.8), S.120 と類似した着想である。なお、同論文では刑罰や責任の時代に応じた意味内容の変化の可能性を根拠に、AI の刑事責任の可能性は否定されないとする。

[173] 稲谷龍彦「Society 5.0 における新しいガバナンスシステムとサンクションの役割（下）」法律時報 94 巻 4 号（2021 年）118 頁（品田発言）。

[174] 稲谷・前掲（注 173）118 頁以下（稲谷発言）。

[175] この指摘は川口・前掲（注 123）112 頁にある。

第４項　過失責任の再考

以上の検討により、AI そのものへの刑事責任の帰属、不規制による解決、そして DPA を含む厳格責任による解決にはそれぞれ重大な問題点を持つことが確認された。そこで本項では、学習を行う AI 製品をめぐる主体における刑事責任の検討を、①開発製造者、②技術サービスプロバイダ、③国・地方公共団体、④利用者、⑤所有者という 3 つの主体に分類して検討を行うものとする。

その際、議論の立場として、AI の利活用に関して妥協を許さない規制を考案するのではなく、相反する利益のバランスを慎重に考慮する必要があるという前提で検討を行う。例えば、一方で AI の誤作動による被害者は、その被害に対する補償を受けるという利益があり、他方、製造者は、自らが開発資金を提供した AI を販売することで利潤を得るという正当な利益がある。また、社会全体としては、信頼性の高い便利な技術の利活用だけでなく、自動運転車がもたらす恩恵を含めた技術の進歩にも関心がある。このような利害の対立がある以上、たとえば過失の刑事責任については、中間的な解決策を模索することになる。

第 1 目　開発製造者

たとえば、AI を搭載した製品によって生命ないしは身体など法益が侵害された場合、それが当該 AI の欠陥によるものであると認定されたならば、製造者には一定の義務が課される。むろん、その「欠陥」は、製造物責任法に従い、製品流通前における、①設計段階における欠陥（設計上の欠陥）、②製造段階における欠陥（製造上の欠陥）、製品流通後における③利用者に対する指示・警告における欠陥（指示・警告上の欠陥）に分類されよう。とりわけ、③については利用者への指示・警告のみでは十分でないならば、当該製品の回収義務が生じうる。以下ではこの 3 つの類型に分類し[176]、刑事製造物責任に関する検討を行う。その際、製造者側に求められる注意義務の具体的内容をとらえる。

（1）製造者の行為

製造物責任の領域において AI 製品については、積極的作為は、欠陥のあるプログラミング、欠陥構造、誤った部品の組立、早すぎたロボットの市場解禁、安全でない、もしくは欠陥のある製品の市場流通などに認められうる[177]。その一方で、たとえば製品にテスト手続や検査手続、もしくは品質制御を実施しなかったり、指示義務を無視したり、製品監視ないしはリコールを実施しなかったりするような不作為の構成要件を満たす可能性がある。その際、積極的作為と不作為の限界づけについて、どの民事上の欠陥カテゴリーが、製造物責任の枠組において存在するのかは問題としなくてよい。なぜなら、そうした欠陥は原則として

---

[176] 以下では,この分類につき *Günther*, Roboter und rechtliche Verantwortung, Herbert Utz Verlag,2016,S.152 ff を参照する。なお、この論文を紹介するものとして松尾・前掲（注 8）73 頁以下。

[177] *Foeste et al.*,Produkthaftungshandbuch, 3.Auflage, C.H.Beck, 2012, § 81, Rn. 4.

いずれかの態度に基づいて惹起されうるものであるからである[178]。しかしながら、大まかな区別も可能であるように思われる。すなわち、製品の欠陥が初めて生じる前に態度が存在するか否か、もしくは、製品の欠陥が公知のものとなった、ないしはそれが認識可能であった後の態度の問題であるか否かに区別されるべきである[179]。

製品の欠陥が初めて現れた場合、製造者の側には、例えば監視要求や注意要求が存在しうる。この注意要求は、例えば消費生活用製品安全法や、消費生活用製品安全法施行令・施行規則、JIS 規格や ISO 規格、もしくは他の技術領域からの規則を通じて立てられうる。積極的作為も、例えば誤ったプログラミングもしくは製造者による早すぎた市場解禁が存在する場合に考えられるが、このことは現実にはほとんどありえないだろう。やはり、欠陥の周知性によって、公知の危険源、もしくは認識すべき危険源に対応しない、もしくは適切に対応しない場合における不作為が主に問題となろう[180]。

（2）保障人的地位の発生根拠

保障人的地位とそれに基づく保障義務の枠組においては、先行の作為から保障義務が生じ、その際ここでは、義務違反的な作為と義務に従った作為に区別されるべきである（先行行為)。さらに、保障義務は引受を通じても生じる（引受による義務)。ここではとりわけ、監視的保障が特別な意味を持つ。

① 先行行為

先行行為による保障人的地位を認めるには、後行する結果発生についての純粋な態度の原因性によって導かれるのではなく、先行する行為が結果発生に近接するという意味での危険を増加させるに至ったといえなければならない[181]。AI 製品の製造者ならば、市場解禁を通じて危険を創出したといえる場合に結果を防止する義務を負うべきといえる[182]。民法上の製造物責任や製造者責任は、この保障人的地位の基礎としても仕えるものとなる[183]。

その場合、義務違反的な先行行為と法に従った先行行為に区別される。さらに、客観的義務違反的な先行行為がすでに存在している限りで、保障人的地位は先行行為により存在し、製造者は結果を防止する義務を負うということが認められる[184]。それゆえ一般的に承認される法規に対応していない場合に、義務違反的な態度が構成される。また、製造者が AI 製品に関し憂慮すべき試験結果があるにもかかわらず、あえて市場に流通させた場合、結果発生の防止についての義務も存在するといえる[185]。

---

178 *Günther*, a.a.O. (fn.176), S.215
179 *Foeste et al.*, a.a.O. (fn.177), § 81. Rn. 5.
180 *Günther*, a.a.O. (fn.176), S.216
181 *Fischer*, Strafrechtgesetzbuch, 66. Aufl., 2019, § 13, Rn. 47.
182 *Foeste et al.*, a.a.O. (fn.177), § 81, Rn. 9.
183 *Fischer*, a.a.O. (fn.181), § 13, Rn. 71 は先行行為の枠組において人的な非難可能性は問題としない。
184 *Foeste et al.*, a.a.O. (fn.177), §81, Rn.9
185 *Foeste et al.*, a.a.O. (fn.177), § 81, Rn. 10,

問題となるのは、たとえばAIに由来する危険が市場流通後になって初めて認識される場合に、法に従う先行行為が、保障人的地位が先行行為から演繹されるのか否かという問題である[186]。例えば、素材の技術や試験技術がさらに発展したことや、AIシステムが自立してさらに発展し、そこで望まれた特性を示さないことが考えられる。ここで問題なのは、製品がもはや製造者の領域に存在しないということである。違法な先行行為に固執しようとする場合、そのような事例では、製造者の負責を根拠付けることができない。そうすると、製造者には、すでに市場に存在する製品にもはや干渉することはせず、その結果、保障人的地位が存在しないということになる。もっとも、製造者が危険な製品を、その危険性を知っていたにもかかわらず、引き続き販売する場合、積極的作為に結びつく刑事責任が考えられる。

AI製品のような複雑なシステムの製造者のみが、自らの製品の危険性を探知する可能性があることが疑わしいことを考慮すれば、製品がもはや製造者の領域に存在しないという理由で製造者の負責を根拠付けられないとすることは説得的なものではなかろう。製品の安全性を信頼し、危険性を知らない利用者は、無防備な状態で危険にさらされることになるだろう。要するに、製造者に義務付けられ、その後認識された危険性を警告させる義務か、もしくはリコールを開始させる義務という製品監視義務が問題となる[187]。その場合、本質的には再び危険源の存在が問題となるため、製品監視義務は、先行行為による保障人的地位の特別事例とみなされる[188]。したがって、その時点ですでにある程度の欠陥が存在していたということになるので、適法な市場流通ではないということになる。結論として、危険源が主観的に義務違反的でなく創出されたか否かは注目すべきではないということになる[189]。

② 引受による義務

製品製造は、分業的なプロセスが取られることが多い。企業の所有者や、企業を代表する権限を与えられた会社組織のみが義務を引き受けることは現実的ではない。そのため、一定の範囲で他者に義務を委任させることができ、その引受で保障人的地位が生じうる[190]。その義務はその人物の活動領域に限定されているので、製品監視に従事する者が、設計上の義務や他の企業分野による義務を履行していないと非難されることはない[191]。しかしながら、委任された義務を引き受ける人物の選出、監督、指導する義務が存在するため、職務を委任したからといって、会社保有者や代表者はその責任から解放されるわけではない。具体的には、会社の組織構造に応じ、場合によっては中間階層を経由して、適切な選択、監督、指導を施さなければならず、これらの義務は権限ある従業員が負うべきであろう[192]。

---

186 Vgl. *Fischer*, a.a.O. (fn.181), § 13,Rn. 50 ff.

187 *Günther*,a.a.O.(fn.176), S.217 この内容については後で議論を行う。

188 *Günther*,a.a.O.(fn.176), S.218.

189 *Kuhlen,* Strafhaftung bei unterlassenem Rückruf gesundheitsgefahrender Produkte - Zugleich Anmerkung zum Urteil des BGH vom 6. 7. 1990-2 StR 549/89 (NStZ 1990. 588), NStZ 1990. 566 (569).

190 *Günther,* a.a.O.(fn.176), S.218.

191 *Foeste et al.*,a.a.O. (fn.177), § 81, Rn. 17.

192 *Kuhlen,* Haftung für Sorgfaltspflichtverletzungen in Unternehmen bei der Produktion von Gütern, in: *Hilrgendorf*

（3）保障義務（作為義務）の発生根拠―製品回収義務との関係

製品流通後に製造者が上記における、指示・警告上の義務を履行したとしても、製造物に由来する被害のためにリコールが利用者から報告され続けた場合、製造者は必要に応じてさらなる被害を防止するために当該製品を回収することが求められうる。この製品回収を懈怠したために、製造者側の責任者が刑事責任を問われた事例がドイツ及び日本において存在する。本項では、それらの事例について学説の見解も踏まえながら検討を行う。

１．ドイツにおける刑事製造物責任の事例

①皮革スプレー事件[193]（BGH St.37 106(1990)）の概要

事案の概要は以下の通りである。1980 年の 9 月末以来、W 社が製造し、S、E 社を通じて販売していた皮革用スプレーの使用に伴い、呼吸困難、咳、吐き気、悪寒、発熱といった健康障害が生じたという被害報告が相次いで寄せられた。被害者の多くは医師の診察を必要とし、中には肺浮腫で生命の危険に晒され、集中治療室に入院した被害者もいた。診断の結果はほとんどの者が肺水腫であった。最初の被害報告の後、返却されたスプレー缶について会社内部で研究がなされたが、原因は解明されなかった。そこで被告人らの製造会社で内部調査が行われ、当該スプレー缶の原料についての調査及び変更を施して販売を続けたものの、いずれの対策も功を奏することなく、被害の報告はその後も続いた。その後、後には E 社のスプレーに関しても被害が報告されるようになった。

そこで、1981 年 4 月中頃に、E 社のスプレーについて製造及び販売停止の措置がとられたが、会社によってその原因を突き止められなかったため、この措置は数日後に撤回された。

その後、1981 年 5 月 12 日に臨時取締役会が開かれ、W 社の取締役である被告人 S 及び Sch も出席した。会社の中央研究所所長(Dr.B）がこの会議で、これまでの検査によれば、スプレーの有毒な性質を示す原因物質を認める根拠がないため、製品を回収する理由は存在しないと述べ、さらに、外部機関に調査を委託し、全てのスプレー缶に警告の表示をし、既になされている表示を改善すること等を提案した。取締役会はこの提案に賛成し、販売中

(Hrsg.), Aktuelle Herausforderungen des chinesischen und deutschen Strafrechts, Mohr Siebeck 2015, S. 210.

[193] この事例を紹介する日本語文献として、稲垣悠一『欠陥製品に関する刑事過失責任と不作為犯論』（専修大学出版、2014 年）9 頁、岩間康夫「刑法上の製造物責任と先行行為に基づく保障人的義務―近時のドイツにおける判例及び学説から」愛媛法学会雑誌 18 巻 4 号（1991 年）41 頁以下、岩間康夫「製造物責任の事例における取締役の刑事責任―集団的決定に関与した者の答責―」愛媛法学会雑誌 22 巻 1 号（1995 年）45 頁以下、岩間康夫『製造物責任と不作為犯』（成文堂、2010 年）5 頁以下、岩間康夫「刑事製造物責任の諸論点―とりわけ回収義務の根拠に関するドイツの議論について―」刑事法ジャーナル 37 号（2013 年）4 頁、北川佳世子「製造物責任をめぐる刑法上の問題点―ドイツ連邦通常裁判所の皮革用スプレー判決をめぐる議論を手掛かりに―」早稲田法学 71 巻 2 号（1996 年）171 頁以下、堀内捷三「製造物の欠陥と刑事責任―その序論的考察―」研修 546 号（1993 年）3 頁以下、ヴァルター・ペロン（高橋則夫訳）「刑法における製造物責任―ドイツ連邦通常裁判所「皮革用スプレー判決』をめぐって―」東洋大学比較法 31 号（1994 年）1 頁以下、松宮孝明『過失犯論の現代的課題』（成文堂、2004 年）21 頁以下など。

止・回収等は、以後の調査により「真の製造物の瑕疵」もしくは「証明可能な消費者へのリスク」が明らかになるまで行なわないことについて一致した。この会議の決定は、その後、S 社の取締役 W と E 社の取締役 D に報告され、両人はそれぞれの会社についてその決定に従った。しかしながら、その後も当該スプレーによる被害は続出し、依然として原因となる物質を突き止めることはできなかった。但し、スプレー缶に表示される注意書きの内容が強化・改善された。そうするうちに、1983 年 9 月 20 日、W 社は連邦保健局及び連邦少年家族保健省の介入により、販売中止及び回収の措置をとったが、回収された製品に含まれる製剤の使用を完全に中止したわけではなかった。この 1981 年 5 月の臨時取締役会をはさむ 2 つの健康被害について W 社の取締役であった S、 Sch、研究所長であった Dr. B、そして販売子会社の取締役であった W 及び D は、1981 年 2 月 14 日から 1981 年 5 月 12 日の臨時取締役会前の健康被害につき過失傷害罪、その後の健康被害につき故意の危険傷害罪で起訴された。LG Mainz で言い渡された原判決によると、被告人 S、Sch、W、及び D に対し、1981 年 2 月 14 日以降の 4 件の被害について過失傷害の併合罪を、また、上記の 1981 年 5 月 12 日特別取締役会の後に発生した 38 件の被害については危険傷害罪を認め、後者の一連の傷害行為のみを、観念的競合として処理した。また Dr.B については、臨時取締役会で不十分な情報提供と助言を行ったために、故意危険傷害罪の幇助を認定した。宣告刑について、S と Sch は、過失傷害につき罰金刑、危険傷害につき 1 年 6 月の自由刑の併科が、W は過失傷害につき罰金刑、危険傷害につき 1 年の自由刑が、D については過失傷害および危険傷害につき罰金刑が、Dr.B については危険傷害の幇助として罰金刑が言い渡された。なお、自由刑については執行猶予付きであった。この有罪判決に関し被告人らは上告した。

②皮革スプレー事件の連邦通常裁判所判決

当該事件の判決要旨は以下のようなものであった[194]。

(1) 侵害を引き起こした物質が何であるのかは不明であるとしても、他の侵害原因の存在が考慮されない場合には、製品の性質とその利用者の健康侵害との原因連関を認めることに法的な誤りはない。

(2) 定められた用法に従って製品を使用した使用者に対して―彼らの正当な予期に反して―健康被害を生じうる危険を基礎づける性質をもつ製品を、製造者および販売業者として市場に流通させた者には、侵害回避が義務づけられる。この義務に有責に従わない者は、それによって生じた侵害に対して、刑法上、不作為によって遂行された身体傷害という点で負責される。

(3) 製造者および販売業者の保障人的地位から、すでに販売されている健康に対して危険な製品を回収する義務が生じる。

(4) 有限会社の複数の取締役が共同して回収命令について決定すべき場合には、取締役全員に、回収命令を決定するために、自己に可能であり、また期待される全てのことを行う義務

194 BGH St.37,106.

がある。
(5)有限会社の取締役が一致して必要な回収をなさないことを決定した場合には、彼らは共同正犯として、不作為から生じた侵害について負責される。
(6) 協同権限(Mitwirkungskompetenz) があるにもかかわらず、必要な回収の決議を実現するのに関与しなかった取締役らは皆、当該措置を怠ったことについての原因を設定したものであるから、仮にある取締役が回収決議を行うよう要求しても他の取締役の反対によって失敗に終わったであろう場合であっても、彼の刑法上の負責は基礎づけられる。
(7) 同一の行為命令の侵害から相次いで複数の侵害が生じた場合には、総じて一個の不作為の所為が存在する。

以下では、主に回収義務の根拠づける上記（２）及び（３）について検討を行う。

２．回収義務の発生根拠に関するドイツの学説

この判例をめぐっては、以下のような学説の展開が見られる。当初提唱されていた回収義務否定説とは、製造者・販売者は製品を販売によって自己の支配から手放して以降は、その製品から発生する危険を阻止する義務をもはや負わないとするものである[195]。

その後に展開されたものとして、皮革スプレー事件原審の見解でもある法令根拠説は、民法上の社会安全義務（Verkehrssicherungspflicht）というこの種の健康被害に関する危険における民法上の製造物責任で認められている製品監視義務から導き出す。BGH は少なくともこの義務の無条件の導入には疑義を示している。この説を支持する Hilgendorf は、この社会安全義務とは、危険源の開放及び操業によって利益を得、危険の操縦に関する独占的地位を有する者にはこれらの危険について責任を負うことが正当に期待されうるという考察から導かれるとする[196]。

BGH の見解を踏襲する学説としては、先行行為説[197]が存在する。これは先行行為に基づく保障人的義務として回収義務を根拠づけるものであるが、先行行為の義務違反性を求める見解と求めない見解に峻別される。

先行行為の義務違反性を求めない義務違反不要説とは、事後にようやく製品の危険性が客観的に認識し得るようになった場合にでも、先にそのような製品を無過失で製造・販売した点になお回収義務の発生根拠を求めるにあたり、先行行為の限定を、義務違反性要件に代えて、「日常行為と比べて危険を高める先行行為」に求める[198]。Kuhlen は、「今日の社会にお

---

195 *Schünemann*, Unternehmenskriminarität und Strafrecht, 1979. S.99.
196 *Hilgendorf*, Strafrechtliche Produzentenhaftung in der Risikogesellschaft, 1993. S. 14l
197 *Kuhlen*, a.a.O. (fn.192), S. 568f.; *Jakobs*, Die Ingerenz in der Rechtsprechung des Bundesgerichtshofs. in: *Roxin/Widmaier* (Hrsg.).50 Jahre Bundesgerichtshof-Festgabe aus der Wissenschaft Bd. IV. 2000. S. 42f.: *Hoyer*. Die traditionelle Strafrechtsdogmatik vor neuen Herausforderungen: Probleme der strafrechtlichen Produkthaltung. GA 1996, S. 160, 176: ペロン・前掲（注 193）11 頁以下。
198 *Meier*, Verbraucherschutz durch Strafrecht? Überlegungen zur strafrechtlichen Produkthaftung nach der "Lederspray"- Entscheidung des BGH, NJW 1992 S.3196

いては、製品の製造及び販売は、当該製品の不相当な危険性が既に販売時に認識し得たか否かにかかわらず、製造業者及び販売業者の保障人的地位を生ぜしめる危険な行為」[199]であるから、この危険な行為に回収義務の発生根拠を見いだすのであると主張する。Perron も、適法ではあるが、客観的に危険な事前行為から保障人的義務が生じることを認め、「経済的企業が、潜在的に危険な製品の製造および販売をすることが許されるのは、これらの製品が販売後もその作用について監視および統制に服されるという条件のある場合だけ」であり「事前には認識できなかった危険が事後的に明らかになった場合には、企業は、消費者を十分保護し、やむを得ない場合には、製品を回収しなければな」らない[200]という。しかし、事後的に判明した危険をも考慮に入れようとするあまりに、先行行為を理由とする不作為犯処罰を無制限に広げてしまう可能性があるとして批判される[201]。

先行行為に義務違反性を求める義務違反必要説は BGH の見解である。ここで、Brammsen は、「刑法上の義務違反性は、刑法において規制化されることによって規定される」[202]のであるから、本判決のように他の法律違反を挙げても解決できるものではなく、「義務を課された者の人格とその者の具体的な義務範囲とに向けられた人的不法論を基礎に、法益を危殆化する態度の義務違反性は常に行為不法のみから規定されうる」[203]ものであるため、「先行行為に基づく（保障人的）義務が発生するのは、少なくとも客観的に過失によってなされた違法な先行行為に限られるのである。つまり、先行する作為によって引き起こされ、高められた具体的法益についての危険は、常に許されないものであり、客観的に予見可能であり、かつ、注意に適合する態度によって回避し得たものでなければなら」ず、「このような刑法上の義務違反性の最低条件が守られなければ、必然的に刑法上の態度秩序が状態保全秩序へと転化されるという負責の拡大へと至」り、「規範の名宛人の決定の自由や影響から引き離されてしまう」[204]と主張する。この学説については、その先行行為者が誰であるかを特定するかという問題[205]や、たとえ製造企業内の役職者に何らかの「危険な」先行行為を認めえたとしても、人事異動等で当該製品を一切管轄しない部署に転属することになった後、あるいはこの役職者が企業を退職した後にまで回収義務を課すことは、実態にそぐわない解決である[206]と批判される。

Brammsen は危険源支配説[207]を提唱する。これは、危険源を支配した点に作為義務の発生根拠を求める立場であり、不真正不作為犯が、作為犯と同様に刑法上の責任を負うのは、行

---

199 *Kuhlen*, NJW 1990, S.568 f.

200 ペロン・前掲（注 193）11-12 頁。

201 北川・前掲（注 193）193 頁参照。

202 *Brammsen*, Strafrechtliche Rückrufpflichten bei fehlerhaften Produkten?, GA 1993, 97, S. 105.

203 *Brammsen*, a.a.O. (fn.202), S.106.

204 *Brammsen*, a.a.O. (fn.202), S.109.

205 岩間康夫「刑事製造物責任の諸論点―とりわけ回収義務の根拠に関するドイツの議論について―」刑事法ジャーナル 37 巻（2013 年）8 頁。

206 岩間・前掲（注 193）96 頁。

207 *Brammsen*, a.a.O. (fn.202), S.108.

為時に結果へと向かう因果の流れを支配することが可能で、結果を防止することができたにもかかわらずそうしなかったからだという考え方に基づき、危険源を支配する者は、そこから法益侵害結果が生じないよう、その危険源を監視する義務が生じる（危険源監視義務）。危険物を製造・販売する者も、自らが作り出し、またその支配領域内にある危険物から法益侵害結果が生じないよう監視する義務があるというものである。もっとも、一度市場流通に置かれた製品の監視には限界があることを留意しなければならない。この学説においては、さらに当該監視の程度の明確な基準が必要とされる[208]。

それ以外にも、情報集中説という、製造者側に当該製造物の欠陥に関する情報のほとんどが集中するという実態に基づき、製造者（あるいは販売者）の製品回収義務を根拠づけようと試みる見解も存在する[209]。これは、製造・販売者が販売後に当該製造物の健康への危険性について認識する場合、彼は消費者に対する情報提供の拒絶によって、消費者による危険の対処に関する任意の選択（危険を受容するか、回避するか、もしくは緩和させるか）を不可能にしてしまうという現象を、製造販売者の情報独占による消費者の処分権に反した形成支配あるいは「知識―物的支配」と呼び、それが従来型の占有による物的支配の存否とは無関係に存在するという。

さらに、引受擬制説[210]という、消費者に対する保護の引き受けの観点を援用して欠陥製造物の回収義務を説明する立場も存在する。これは、製造物資任における製造者・販売者の義務内容を、製品の現実の回収ではなく、製品使用に関する警告へと修正した上で、この警告義務の本質は製造者が消費者に対し、まさに警告を行うことを約束し、消費者が製造物の危険がないようにする配慮を製造者の手に委ねるという形で信頼するというところにあるものである。この学説の特徴は、不真正不作為犯成立の統一的基準として掲げた「結果の原因に対する支配」の下位事例としての「被害者の脆弱性に対する支配（保護的保障の根拠）」を製造・販売企業に認めることにある。なお、Roxin の見解では、商品の危険や欠陥に関する情報が製造者側にしかなく、消費者は製造者に欠陥の存在やその原因等の解明や対策について製造者を信頼し、製造者は相応する範囲で消費者の保護を引き受けなければならず、その場合の製造者・販売者の義務内容は警告には限定されないとする[211]。

以上、皮革スプレー判決をめぐって様々な学説が展開されたが、回収義務否定説及び法令根拠説を除く学説は、製造者側の処罰を求めるが故に、従来の作為義務の内容につき、例えば「排他的支配」や「情報の集中」、「引受擬制」という法律上の基準ではない幻の基準と「先行行為」とが組み合わせることで、義務付けの限界が不明確になるおそれがあると批判される[212]。むしろ、作為義務の内容を確定するための明確な基準として、法律上の基準はそれを

---

208 後述する「パロマ湯沸器事件」判決も参照。

209 *Bottke*, Krankmachende Bauprodukte-Produkthaftung aus zivil- und strafrechtlicher Sicht unter besonderer Berücksichtigung krankmachender Gebaude (Sick Building Syndrom) Teil 2., ZfBR 1991, S.237 f.

210 *Schünemann*, Unternehmenskriminalität.in: *Roxin/Widmaier* a.a.O. (fn.197), S.638 ff

211 *Roxin*, Strafrecht Allgemeiner Teilband II, 2003, §32 Rn. 210 ff.

212 岩間・前掲（注 205）10 頁。

導く一つのファクターとなるため、存在する限りで法律上の基準を設けるのが一つの解決方法となりうるのである。

３．日本の刑事製造物責任における作為義務の根拠の学説

回収義務の不履行は、過失不作為犯の注意義務違反とされる。そのため、不作為犯の要件としてみなされる「保障人的地位」ならびに「保障義務（ないしは作為義務）」の発生根拠についての日本の議論を概観する。

まず、通説的見解である形式的三分説とは、法令・契約・先行行為から保障人的地位を求めるものである。むろん、これらは直ちに作為義務の発生の根拠とはならないという批判がある。例えば、民法上の義務（扶養義務など）が直ちに刑法上問題となる作為義務を発生させるわけではない。そこでドイツの議論を踏まえる形で日本でも様々な学説が展開された。その一つが先行行為説である。この学説は、不作為を作為と同視するには、不作為以前に法益侵害に向かう因果の流れを設定したことが必要であり因果の流れを設定する（故意または過失による）先行行為が作為義務を発生させるとする見解である[213]。ただし、この見解では、販売後に隠れた欠陥が発覚した場合、人員の異動等配置転換がある場合に作為義務を根拠付けるのが困難とされる。次に、事実上の引受説とは不作為者が法益保護を事実上引き受けている場合に作為義務が発生するとする見解である[214]。より具体的には、①法益の維持・存続を図る行為の開始／存在、②そのような行為の反復・継続性、③法益の保護についての排他性の確保を考慮し、結果発生の有無が具体的に不作為者に依存するという事実上の引受行為が生じているか否かによって判断すべきとする。もっとも、この見解では危険な欠陥製品が自己の支配領域内にない場合は、リコールすべき作為義務（回収義務）は生じないとするところ、後述する薬害エイズ（厚生省ルート）事件やパロマガス湯沸器事件などの判例を説明できないことに注意しなければならない。

次に、排他的支配領域性説という、不作為者が自己の意思で排他的支配を獲得したこと、また、自己の意思によらない場合には、事実上結果を支配する地位が生じたことについて規範的要素を考慮することにより、作為義務が発生するとする見解がある[215]。この見解における「支配」概念については、事実的支配では上記の事実上の引受説と同様の問題を生じうるので、規範的支配で足りる。また、「排他性」概念については、文字通りの排他性を貫徹すると、薬害エイズ（厚生省ルート）事件やパロマガス湯沸器事件における二次的責任を負った被告人の責任について説明できなくなる。そこで、「情報の掌握」という観点から、排他性を基礎づけるのが有効ではないか[216]とする見解もある。これによると、独占的に情報を掌

---

213 日高義博『不真正不作為犯の理論（第 2 版）』（慶應通信、1983 年）154 頁。

214 堀内・前掲（注 193）8 頁。

215 西田典之（橋爪隆補訂）『刑法総論（第 3 版）』（弘文堂、2019 年）132 頁。

216 岡部雅人「刑事製造物責任における回収義務の発生根拠―わが国の議論状況をめぐって―」刑事法ジャーナル 37 号（2013 年）15 頁。

握することにより、規範的支配を危険源に対して及ぼしていたといえるような、結果回避措置をなしうるだけの地位・職責・権限を有する者に作為義務が認められるとする[217]。しかし、この理論では、パロマ湯沸器事故や三菱自動車車輪脱落事故は説明できても、薬害エイズ（厚生省ルート）事件の被告人は、国産原料であると偽って販売を続けた「ミドリ十字」関係者がいる中で、その情報を知っていたという証拠が認定されていないにもかかわらず、同事件での非加熱製剤の危険性をコントロールする「排他的」支配があるとはいえないと批判される[218]。

さらに効率性説という、結果に対して因果性を有する行為者が複数存在する場合には、「結果回避措置を最も効率的に為しうる主体のみが保障人的地位に該当」するとする見解がある[219]。その内容として、結果回避措置との関係において、企業内部の者が①製品に関する危険情報を掌握していること、とりわけ回収義務の場合はそれに加えて、流通ルートの把握の程度も考慮する、②その者が、当該結果回避措置をなすことについての意思決定をなしうる地位にあることを要件とする[220]。ただし、薬害エイズ（厚生省ルート）事件における被告人やパロマガス湯沸器事件などの被告人はそのような地位であったのかは疑問である。

その他、社会的期待説という、社会の「期待」こそが「主体」を選び出す基準として着目されるべきであり、「法益が特に危険に晒されやすく、その要保護性と不作為者が特別の関係に立つと見られる場合、又は、危険源が日常生活上のものを超える危険を潜在的に含んでおり、その危険性と不作為者が特別の関係に立つと解される場合には、不作為者に刑法上の保障人としての『主体』性（身分）を肯定しうる」とする学説がある[221]。しかし、この見解に対しては、社会的期待とは結局、社会に現に存在する道徳的規範の言い換えにすぎず、このような見解では刑法上の義務と道徳上の義務の区別は質的なものではなく量的なものであるからして、保障人的地位の範囲を不明確にするおそれがあると批判される[222]。

最後に、機能的二分説[223]という、作為の機能に着目して、その発生根拠を大別して２つの場合に分類する見解がある。その一方は、当該法益を保護すべき関係に立つ場合の義務の作為義務の発生根拠（保護的義務）、そしてもう一方は、危険源を管理・監督すべき義務が認められる場合の作為義務の発生根拠（危険源監視義務）とする[224]。まず、保護的義務につい

---

217 岡部・前掲（注 216）16 頁。

218 松宮孝明「薬害エイズ事件厚生省ルート事件最高裁決定」医事法学 24 号（2009 年）162 頁。

219 鎮目征樹「刑事製造物責任における不作為犯論の意義と展開」本郷法政紀要 8 号（1999 年)）355 頁。

220 鎮目・前掲（注 219）369 頁。

221 木村亀二「不作為犯における作為義務」木村亀二『刑法解釈の諸問題第一巻』（有斐閣、1939 年）248 頁、塩見淳「瑕疵ある製造物を回収する義務について」刑法雑誌 42 巻 3 号（2003 年）370-371 頁など。

222 佐伯仁志「保防人的地位の発生根拠について」内藤謙ほか編『香川達夫博士古稀祝賀論文 刑事法学の課題と展望』（成文堂、1996 年）101 頁。

223 山中敬一『刑法総論（第３版)』（成文堂、2015 年）244 頁。

224 この着想は不作為犯の文脈における正犯の背後者たる共犯を処罰するための根拠となりうることにも留意すべきである。

ては、現実的に危険な状況に陥っている行為客体が、保護を要する状況にあることを作為義務の前提とする。継続的な社会的身分や地位に基づく関係という規範的保護関係に基づく作為義務、不作為者と被害者との間の合意により保護機能を引き受けるべき関係が設定されている場合、ないしは黙示の合意、すなわち危険共同体の場合における任意的・制度的保護関係に基づく作為義務、被害者の法益が、不作為者の先行する法益維持行為によって機能的に支えられており、それがなければ法益が失われていた場合のように、法益が機能的に特定人の作為に依存しているという機能的保護関係に基づく作為義務に分類される[225]。それに対して危険源監視義務とは、危険源による危険が被害者に対して現実的に危険な状況を創出していることを作為義務の発生根拠とする[226]。この場合、たとえその危険が現実化していなくても、一定の結果発生をもたらす危険源が存在し、当該不作為者がそれに対する管理・監督責任を負い、それを履行していないことが前提となる。そこには、危険な物・設備に関する管理義務に基づく作為義務、第三者の危険行為に関する監督義務に基づく作為義務、不可罰の先行危険行為創出行為に基づく作為義務に分類される。刑事製造物責任の文脈では、専ら危険源監視義務が問題となる。すなわち、製品流通後に判明した欠陥によって生じた被害について製造者に回収を義務づけるには、その製品が一定の結果発生をもたらす危険源であるとしこれを管理（監視）する義務を負うとすれば、製品回収もまたこの範疇とすることが可能であるから、よって製造者に対して刑法上の「回収」義務を肯定しうる。

以上、刑事製造物責任をめぐる学説を概観してきたが、これら学説は製造者に対する「回収」義務を根拠づけることに重きが置かれていることに注意すべきである。しかし、製造者が刑法上の結果を回避できる措置としては、回収だけでなく消費者に対する製品の警告に関する情報提供[227]や指示を行うことでも達成されないかという指摘がある[228]。私見としても、製品欠陥に起因する結果発生に関し、製造者には直接的に刑事責任を検討するのではなく、行政法規など一定の法規範の中で製造者に一定の義務付けを行い、その義務違反に対して罰則を設けるというスキームが妥当ではないかと考える[229]。そこでは民法上の製造物責任の構造が参考となる。

（4）作為義務の類型

分業体制を取る製造者内部においては、管轄の異なる主体が複数存在するため、製造物責任の枠組における注意義務は、製造者の活動に応じて区別されるべきである。これら主体に対しどのような義務が課されるのかを、（3）で示唆した民法上の製造物責任のスキームを

225 山中・前掲（注 223）244 頁以下。

226 山中・前掲（注 223）245 頁以下。

227 岡部雅人「刑事製造物責任における『回収義務』について」早稲田大学大学院法研論集 123 号（2007 年）116 頁以下、岡部・前掲（注 216）16 頁。

228 岩間・前掲（注 205）10 頁。

229 この見解は、甲斐勝則「欠陥製品の製造・販売と刑事過失」斉藤豊治・日高義博・甲斐勝則・大塚裕史編『神山敏夫先生古稀祝賀論文集』（成文堂、2006 年）176 頁も参照。

援用しながら考察する[230]。すなわち、①設計・製造段階、②指示・警告段階、③製品監視段階の分類[231]にしたがって先行研究のスキームに従い検討を行う。

① 設計・製造段階

この段階で製造者に対して求められる義務として、製品がその設計の結果求められる ISO 基準、JIS 規格などの安全基準に適するようにする設計上の義務と、個別の製品が設計に従って安全基準を満たすようにする製造上の義務が挙げられる。

これに関する先例として、森永ドライミルク事件差戻審第一審（徳島地判昭和 48 年 11 月 28 日判時 721 号 7 頁）では、薬事法および工場に領置していた書籍「註解第六改正日本薬局方」、「薬の原理とその応用最新薬理学」の抜粋複写写真[232]を注意基準としたうえで、製造課長・工場長の注意義務を判断しており、組立式サウナ事件控訴審判決（東京高裁昭和 53 年 3 月 28 日刑集 33 巻 7 号 748 頁）でも、判文中ではサウナの構造については専門家（塚本孝一）の鑑定を、無焔着火の原理については、専門家の鑑定ならびに建築学会発行「建築学会論文集（第一六号）」（1940 年 2 月発行）所収の浜田稔・平山嵩共著「木材の加熱による出火の可能性に就て」、「建築学大系」全 21 巻（1956 年初版発行）所収の「建築防火論」などの論文を用いて注意義務を認定する。まとめると、問題となる製品が満たすべき基準を充足しているか否かは、その当時の科学技術の水準に応じた規格基準ならびに専門知識により判断される。その際、上記のものがどれほど普及しているかに着目して、製造者はこれらに関する情報の収集に努める義務を負うべきである。

② 指示・警告段階

この段階では、製造者が欠陥なく製品を製造したが、その利用において生じる危険を注意喚起しなかった場合が想定される。具体的には、製造者が必要な警告を行わなかった、ユーザーへの情報提供が不十分であった、既存の危険性を些細なこととして示した場合などが挙げられる。この関連で製造者の指示・警告に関する瑕疵により刑法上の問題を招来した先例としては、厳密には製造物に関するものではないが、渋谷温泉爆発事故事件（最決平成 28 年 5 月 25 日刑集 70 巻 5 号 117 頁）が挙げられる。その決定要旨において、「被告人は、その建設工事を請け負った本件建設会社におけるガス抜き配管設備を含む温泉一次処理施設の設計担当者として、職掌上、同施設の保守管理に関わる設計上の留意事項を施工部門に対

---

230 民事製造物責任上における製造者に対する社会生活上（取引上）の義務と結びつける着想は、*Hilgendorf*, a.a.O. (fn.196), S.141 f.と親和的である。実際、彼が主催する RobotRecht チームの研究論文である *Günther*, a.a.O (fn.176), 233 f.や *Lohmann*, a.a.O. (fn.49), S.163 f.もこのような分類を行い、その各々に従って製造者の刑事責任を検討する。

231 消費者庁消費者安全課『逐条解説 製造物責任法（第 2 版）』（商事法務、2018 年）58-59 頁を参照。なお、このスキームに従って刑事製造物責任を検討するものとして中川・前掲（注 100）23 頁以下。

232 判文では「国家が制定した医薬品の公定書であり、医薬として基礎的に重要性のある代表的医薬品を収載し、その強度、品質及び純度の基準を定めたもの」とする。

して伝達すべき立場にあり、自ら、ガス抜き配管に取り付けられた水抜きバルブの開閉状態について指示を変更し、メタンガスの爆発という危険の発生を防止するために安全管理上重要な意義を有する各ガス抜き配管からの結露水の水抜き作業という新たな管理事項を生じさせた。そして、水抜きバルブに係る指示変更とそれに伴う水抜き作業の意義や必要性について、施工部門に対して的確かつ容易に伝達することができ、それによって上記爆発の危険の発生を回避することができたものであるから、被告人は、水抜き作業の意義や必要性等に関する情報を、本件建設会社の施工担当者を通じ、あるいは自ら直接、本件不動産会社の担当者に対して確実に説明し、メタンガスの爆発事故が発生することを防止すべき業務上の注意義務を負う立場にあったというべきである」としている。ここでは被告人の職掌が（結果回避）義務を根拠づけていることが窺える。

③ 製品監視段階

この段階では、製品が市場に流通した後に、人の生命・身体への侵害が生じるリスクを回避する場合に製造者にはどのような義務が課されるかが問題となる。この点、大多数の消費者の手に渡った製造物を各々管理・監視することは到底不可能であり、仮にこのような義務を製造者に課すのは多大な負担となることは考慮するまでもない。そこで、先例を参照しつつ、どのような場合に製品流通後における結果発生が製造者に帰属されるのか、その条件を捉えていく。

六本木回転扉事故事件（東京地判平成 17 年 9 月 30 日判時 1921 号 154 頁）では、「被告人Ｂは、社内の事故速報を通じ、本件事故前にも自動回転ドアでの複数の事故情報に接し」さらに、「戸先と固定方立との間に児童の体がしばらく挟まれ、その際、児童が頭部挫創等の傷害を負った」事故内容の「報告を受け、かつ、その場で、設計本部としても、大型自動回転ドア『シノレス』の安全対策を検討するよう求められた。したがって、被告人Ｂとしては、遅くともこのころには、シノレスに設置されている危険防止機能が十分なものではなく、そのままの状態でシノレスの使用を続ければ、挟まれ事故が発生して、通行人を死傷の結果に至らしめることを予見することが可能であったと認められる」として、「被告人Ｂは、本件事故回避のために本件シノレスに防止柵を取り付けたり、あるいは、緩衝材の取り替えを実施したりすることが十分に可能であり、結果回避可能性があったことに何ら疑いはなく、前記の進入を防止する義務も、死傷の結果を生じさせない装置を取り付ける義務も共に履行されていなかったといわなければならない」として被告人の職掌および事故情報の集中を手がかりに作為義務を認定している。

また、パロマガス湯沸器事件（東京地判平成 22 年 5 月 11 日判タ 1328 号 241 頁）でも、「被告人甲山は，パロマ両社において，昭和 56 年 3 月以降本件事故までの間，代表取締役社長ないし同会長として，製造販売品の安全確保，事故対応，リコールを含む業務全般を統括し，これらについて事実上の最終決定権限を有していたのであるから，…平成 13 年 1 月 5 日ころから本件事故までの間において，自らないしは被告人乙川等のパロマ両社の関係部

署の担当者らに指示するなどして，上記注意喚起の徹底，点検・回収の措置をとるべき刑法上の注意義務を負う立場にあ」るとして被告人の作為義務を認めている。ここで特筆すべきなのは、直接の事故原因が下請業者による「短絡」であったにもかかわらず製造者側の品質管理部門の最終責任者たる被告人に刑事責任が問われていることである。

さらに、薬害エイズ（ミドリ十字ルート）事件（大阪高判平成 14 年 8 月 21 日判時 1804 号 146 頁）では、「被告人Aは，血液製剤等の医薬品の製造販売等を業とするDの代表取締役社長として，同社の業務全般にわたる重要な案件について協議し決定する機関である常務会と経営会議を主宰し，営業方針等について報告を受けるなど同社の業務全般を統括していたもの，被告人Bは，同社の代表取締役副社長兼研究本部長として，常務会等を構成して同社の意思決定に参画し，被告人Aを補佐して同社の業務全般に関与すると共に，エイズと血液製剤との関わりについての情報収集等の調査を含む医薬品の研究に関する業務を統括していたものであり，いずれも同社の医薬品の製造販売に伴う危険の発生を未然に防止すべき地位にあった」としたように被告人らの職掌が作為義務の検討の際に考慮に入れられている。

最後に、三菱自動車車輪脱落事件（最決平成 24 年 2 月 8 日刑集 66 巻 4 号 200 頁）では事故関係の情報を一手に把握していたことを踏まえつつ、「被告人Yについては，その地位や職責，権限等に照らし，関係部門に徹底した原因調査を行わせ，三菱自工製ハブに強度不足のおそれが残る以上は，被告人Xにその旨報告して，関係会議を開催するなどしてリコール等の改善措置を執り行う手続を進めるよう進言し，また，運輸省担当官の求めに対しては，調査の結果を正確に報告するよう取り計らうなどして，リコール等の改善措置の実施のために必要な措置を採り，強度不足に起因するDハブの輪切り破損事故が更に発生することを防止すべき業務上の注意義務が」、また、「被告人Xについても，その地位や職責，権限等に照らし，被告人Yから更に具体的な報告を徴するなどして，三菱自工製ハブに強度不足のおそれがあることを把握して，同被告人らに対し，徹底した原因調査を行わせるべく指示し，同社製ハブに強度不足のおそれが残る以上は，関係会議を開催するなどしてリコール等の改善措置を実施するための社内手続を進める一方，運輸省担当官の求めに対しては，調査の結果を正確に報告するなどして，リコール等の改善措置の実施のために必要な措置を採り，強度不足に起因するDハブの輪切り破損事故が更に発生することを防止すべき業務上の注意義務があったというべきである」として、これも被告人らの職掌に照らして作為義務を認定している。

以上の先例を通した、製品流通後における製造者の作為義務の認定スキームをまとめると以下の通りとなる。すなわち、①当該製品による事故に関する情報が企業内部の行為者に集中し独占されていること（予見可能性の基礎)、②その上で改善措置等の検討をすべきところ、これを懈怠したこと（結果回避義務違反の基礎①)、③行為者に製品を管理する権限が与えられている、または行為者でない第三者による行為が結果の原因として介在する場合は、行為者がその第三者の対する指揮・監督権限を有していること（結果回避義務の基礎

②)、④その改善措置・回収措置を行わなかった結果、同種の被害が発生したといえること（因果関係の基礎）である。この中で、②を認定するには法的基準ないしは刑罰法規のみならず、行政法規・民事法規も含めた法的期待状況の存在[233]が望ましいとされる。さらに、法的義務のレベルで規定されてはなくとも、欠陥のある製造物による人の生命・身体被害に対して法秩序がどのような考え方に立つのかを認識するもので足りるとする見解[234]も存在するが、依然として規範的判断の域を超えないので、少なくとも法律上で明文化された基準を求める必要があろう。事実、薬害エイズ（ミドリ十字ルート）事件の後、企業倫理の遵守に資する規定が創設された。すなわち、薬機法 68 条の 9（現行）は回収義務を含む薬害防止に関する義務規定の創設であり、その実効性を担保するために同法 69 条で行政による立入検査を実施することが可能となっている[235]。

（5）AI 製品の製造者に課せられる義務内容

（4）で見たように、製造者に対して課せられる注意義務には刑法外の法律上の義務が関連することがある。では、AI 製品の製造者に対して課せられる義務内容とはどのようなものとなるか。以下、先述の分類に従い、①設計・製造上の義務、②指示・警告義務、③製品監視義務に分類して具体的に検討したい。

① 設計・製造上の義務

製造者は、製品構想の枠組で、最新の技術水準を、流通に置いた時点で顧慮せねばならず、規定に従った利用を通じて設定される安全基準を遵守しなければならない[236]。製造者がこの義務を下回る場合に設計上の欠陥が存在する。JIS 規格のような、企業相互間の技術的規範が、下限値として考慮される場合、それらが存在する限りで、その製品が技術水準に対応していることを推定させる[237]。しかし、技術開発がこれらの基準や規則を超えた場合、承認された技術水準に加えて科学技術の水準を考慮するなど、製造者はこれを超える措置を講じなければならないとされる[238]。日本のロボット工学の領域では、すでに ISO 規格が存在する。たとえば、2016 年に公表された、生活用途の支援におけるロボットならびにロボテ

---

233 山中敬一「刑事製造物責任論における作為義務の根拠」関大法学 60 巻 5 号（2011 年）64 頁以下。

234 山中・前掲（注 233）65 頁以下では、「特定製品」（消費生活製品安全法 2 条）において製造者に対して事故原因の調査・製品の回収措置ならびにその他の危害の発生・拡大防止を義務付ける消費生活製品安全法 38 条 1 項や、薬機法 68 条の 9（旧薬事法 77 条の 4）に規定する回収義務、そして行政官庁（厚生労働大臣）に対しては薬機法 69 条の 3 に規定する緊急命令規定、自動車のリコール制度を定める道路車両運送法 63 条の 3 を参照する。

235 ただし、罰則規定は存在しない。

236 土倉澄子『逐条講義 製造物責任法 第 2 版 基本的考え方と裁判例』（勁草書房、2018 年）219 頁以下。

237 土倉・前掲（注 236）223 頁。

238 *Foeste et al.*, a.a.O. (fn.177), § 24, Rn. 47

ィクスデバイスについての ISO 13482 に準拠した JIS B 8445 が挙げられる。ここでは、危殆化とリスク低減のための可能な措置に関する規則を制定し、このようなシステムの制御の堅固性に関する最低要件が定義され、技術システムの「自律性」も考慮されるようになった[239]。 そこでは、以下のように記載されている。

「5.12 誤った自律的判断及び動作による危険源

5.12.1 一般 自律的に判断を下し、動作するよう設計された生活支援ロボットは、間違った判断及び誤った動作が受容できないリスクの原因とならないよう設計しなければならない。

例 1 移動作業型ロボットが間違った飲物をつかみ、一杯の水の代わりにコーヒーを出すとすれば、それは受容可能なリスクであるが、もし割れたカップに入った飲物を出すのであれば、そのリスクは受容できない。

例 2 搭乗型ロボットが、平たん（坦）な地面において急に予期せぬ回避動作をとるとすれば、それは受容可能なリスクであるが、滑りやすい地面の方へ回避動作をとるならば、そのリスクは受容できない。 誤った判断の影響として生じる危害のリスクは、判断の信頼度を上げる（例えば、より良いセンサの使用）か、又は誤った判断の影響を制限する（例えば、使用限界を狭める。）のいずれかによって低減することができる。」

より高い信頼性に努めたり、誤った決定の効果を軽減したりすることで誤った決定から生じるリスクを低減できることができると認められる。上記規定は、まだあまり具体的ではないが、ロボットの分野での新たな展開を取り上げている。なお、古い規格しか存在しない場合や、規格が存在しない、または計画されていない場合にも問題が生じる。その一例としては、人間と協働する産業用ロボットが衝突した際の力や圧力の限界値がある。現在、関連する ISO 10218-1 や ISO10218-2 には、具体的な制限値はいまだ含まれてはいない[240]。そのため、基本的には、製造者が入手できる最新の技術的・科学的知見に従った設計を方向づけなければならない。さらに、新たな規格や規制について常に情報を得るだけでなく、特に規制が存在しない領域では、現在の展開を把握する必要がある。また、規範が古くなればなるほど、可能な限り早く最新の科学技術に基づく規範を照会することが必要となるが[241]、これには、国際科学会議、専門イベント、専門文書などが含まれるものとされる[242]。

製造上の欠陥は、個々の部品が製造過程の枠組において、欠陥のない設計の場合における安全性の水準を達成できない場合に存在しうる。製造上の欠陥の原因は、不注意、機械の摩耗、もしくは製造機器の誤操作など多岐にわたる。このような製造上の欠陥や、回避可能な

239 なおこの規格における「自律性」とは,「人が介入することなく,現在の状態及びセンサ計測に基づいて、意図したタスクを実行する能力」と定義される。

240 なお、日本における労働安全衛生規則 150 条の 4 では,定格出力が 80W を超える産業用ロボットに接触することにより危険が生ずるおそれがあるときは、柵または囲い等を設けること、と規定する。

241 *Foeste et al.*, a.a.O. (fn.177), § 24, Rn. 47.

242 *Foeste et al.*,a.a.O. (fn.177), § 24. Rn. 33

「外れ値」に対応するために、品質管理はそのような欠陥を発見し、これを取り除く製造者の「重要な義務」として存在する[243]。もっとも、品質管理の手段は多岐に亘るため、個別事例に応じてその意義や有用性が異なる。AI 製品の分野では、目視検査、機械的検査、自動測定器による検査などが重要な役割を果たすが、品質管理の手段は科学技術の水準に適合したものでなければならない[244]。例えば、目視検査によって、設計に反する組立、固定不良、もしくはケーブル接続の緩みなどの明らかな欠陥を発見することができる。また、機械検査では、車軸の移動度、部品の耐荷重性、速度、精度、あるいは安全機能を検査することができる。さらに自動測定器による検査は、通常、オートメーション、エレクトロニクス、またはコンピュータの利用を意味し、AI 製品のハードウェア機能をチェックするテストプログラムにより、特にソフトウェアとハードウェアの連携において、目視や機械的なテストの枠組では明らかにならない他の欠陥を検出することができる。例えば、製造領域でのソフトウェアの欠陥は、伝送の不具合や、プロセッサやメモリーモジュールの不良などのハードウェアの不具合によって存在し、システム制御に影響を及ぼすことがある。さらに、自動検査には、純粋な機械的検査や目視検査に比べて、ヒューマンエラーを排除できるという利点がある[245]。このように、製品製造においては種々の管理手法が存在し、その全てを網羅することは困難であるが、少なくとも製造者に対しては「品質管理」に努める義務を策定しても良いのではないかと考える。

② 指示・警告上の義務

製造者は AI・ロボットの危険性を利用者に対して適切な指示を与える、もしくは警告しなければならない。これは特に利用者による正しい使用方法だけでなく、ユーザーによる誤用にも当てはまるが、とりわけ誤用の場合は、製造者の製造物責任と消費者の成熟性との間で適切なバランスを取る必要がある[246]。その指示義務は、顧客の AI に対する期待、AI がもたらす危険の程度、そして個別事例で AI 製品によって危殆化される人物によって具体化されうる。自己学習する AI を搭載したロボットの場合は、特にその指示に目を向けるべきであり、製造者は AI・ロボットの態度について特定の状況下では予見できないため、予防的に指示を通じて起こりうる危険に対処しなければならないだろう。

この指示義務は、AI を備えたシステム自らの学習により発展する能力の場合に問題となる。例えば、危険を警告する義務は、その危険が製造者に認識できる限りにおいてのみ存在する。そのため、遠隔的な危険については必ずしもこのことを製造者に期待することはできない[247]。しかし、何がまだ認識可能で、何がそうでないのかの線引きはどこにあるのかにつ

---

243 *Foeste et al.*,a.a.O. (fn.177), § 24, Rn. 194.
244 *Foeste et al.*,a.a.O. (fn.177), § 24, Rn. 204 ff.
245 *Foeste et al.*,a.a.O. (fn.177), § 24, Rn. 210.
246 *Foeste et al.*,a.a.O. (fn.177), § 82, Rn.33; ここでは民法上の責任枠組における指示義務についての説明が援用される。
247 *Günther,* a.a.O. (fn.176), S.156.

いては、事前の危険調査の領域でも、製造者にとって大きな挑戦となりうる。というのも利用者よりも製造者の方が製品に精通しているであろうから、より発展性のあるシステムの場合には、より高次の基準で指示義務を課すことが正当化されるだろう[248]。これを十分に遵守するためには、製造者はその後の製品の使用状況を考慮して適切な調査を行い、明らかになった危険性を精確に分析しなければならない。また、製品流通後の指示義務・警告義務については、例えば、製品に内在する危険性が、科学技術の新たな知見の結果として後になって判明した場合、製品が新たな事情の発生によりリスクをもたらす場合、あるいは科学技術の新たな知見の結果として製品の危険性にから法益が保護される場合などがこれに該当する[249]。製造者には必要に応じて、事後的に説明を行う義務がある。さらに、指示義務の出発点として考えられるのは、特に学習能力を有する AI システムにおける「可読性」[250]の構想である。その背景にある理念は、AI 製品が利用者にとって「読むことができる」存在になるということにある。従来、指示は、マニュアル、説明書、文書などによって果たされてきたが、あらゆる種類の指示が常に同じように適しているわけではない。したがって、ある状況下でシステムが自ら指示を出し、製造者がそれによって指示義務の一部を果たしうることは少なくとも考えられる。

いずれにしても、上記で取り上げた義務そのものは法律上の義務として明文化されているわけではない。その個々の内容を網羅することは設計・製品上の義務と同様に困難ではあるものの、少なくとも製造者は「利用者に対して適切な指示・警告を与える義務」を行政法規ないしは法的拘束力は持たなくともガイドラインのレベルで明文化すべきであろう。なぜなら、回収義務の根拠の議論でも見たように、およそ法律上の義務にないものを根拠にして刑法上の注意義務違反を認定するということは過失の範囲を無制限に拡張することになりかねないからである。

③　製品監視義務

製造者の義務は、一見して欠陥がないように見える製品を市場に流通させたという事実だけでは消尽しないことは先例のとおりである。そこで考慮されうる科学技術の水準は、市場流通後に発生するリスクを特定するための尺度として、また、その危険を防止するための措置の尺度として機能する[251]。ここでは、データ保護に配慮しつつ、新たな技術を利用することができるが、以下の事項が考慮される[252]。すなわち、エンドユーザーが記入するフィー

---

248 *Günther,* a.a.O. (fn.176), S.156.

249 *Günther,* a.a.O. (fn.176), S.157 ff.

250 *Günther/Münch,* Legal lssues of Making a Robot “Readable”, in: Workshop on Robot Feedback in Human-Robot Interaction: How to Make a Robot,,Readable" for a Human lnteraction Partner, 21st IEEE International Symposium on Robot and Human lnteractive Communication, 2012.

251 *Günther*, a.a.O. (fn.176), S.159.

252 *Moliroris/Klindt,*Produkthaftung und Produktssicherheit – Ein aktueller Rechtsprechungsüberblick, NJW 2008, S.1203 によると、チャットルームもしくは YouTube 動画などが製造者の市場監視ツールとして述べられている。

ドバックフォームから、AI がデータネットワークを介して製造者に自動的に送信するステータス信号に至るまでである。これらのステータス信号は一般的に、システムの状態に関する匿名の情報を提供したり、障害が発生したりする場合には、個別のメッセージで製造者に通知することができる。製造者側が受け取った情報は、直ちに処理・分析しなければならない一方で、AI 製品に欠陥があることを知る限りで、危険回避措置を講じなければならない[253]。AI 製品がネットワーク化されていれば、情報をその製品に送ることも可能であろう。そのような情報としては、警告指示、リコール要求、またはソフトウェアのアップデートを自動的にインストールするという形で考えられる。さらに、ソフトウェアの場合、製造者は定期的なアップデートの必要性や、特定のリスクにも注意を払わなければならないだろう。とりわけ AI の領域では、システムの挙動が製造者にとって予測不可能であるものもありうるため、製品監視義務は非常に重要な意味を持つ。とりわけ、製品を詳細に観察することでしか認識できないリスクもありうるので、学習能力のあるシステムにおける製品監視、それに伴う措置の義務付けはこの意味で必要だと思われる。

（6）義務違反と結果との因果関係：不作為の因果関係

刑事製造物責任においては、行為と結果の間にどのような科学的連関があるのかを疑いの余地なく立証できるとは限らない。例えば、先述した「皮革スプレー判決」では「因果関係が存在することのみが確定している限り、科学的知見によれば、最終的に何が原因で損害が発生したのか」を立証する必要はないという形で締めくくられている[254]。これは、AI・ロボットの領域では非常に重要なことであり、開発者、製造者、外部の専門家のいずれもが、しばしばその因果関係を科学的に説明できない、少なくとも意見が一致しないことがある。ここで、いわゆる一般的な因果関係で十分だとすると、つまり、ある製品が何らかの形で構成要件的結果をもたらしたと仮定すると[255]、その結果に対して理解できる他の説明が除外されうることが前提条件となるが、この「消去法」では、科学的に疑問のある論証をするリスクがあることも否めない。すなわち、裁判官は裁判で、他のすべての損害原因が存在しないため、製品の欠陥のみが損害について答責的でありうるという知見を得なければならない[256]。これには 2 つの問題がある。一つ目は、裁判官は自らが知っている予備的原因しか排除できないため、客観的な不明確性が必ず残ることである。二つ目として、予備的原因を排除することを重要視しすぎると、「疑わしきは被告人の利益に」原則の妥当性が否定されてしまうことが問題となりうる[257]。少なくとも、認められた科学的知見を看過したり、科学的に承認されていない理論を適用したりするような手続であってはならないといえる。因果関係との関連では、製造工程における分業から生じる問題、すなわち、複数人の異なる誤っ

253 *Günther*, a.a.O. (fn.176), S.160
254 *Beulke/Bachmann,* Die „Lederspray-Entscheidung“ -BGH St 37, 106, JuS 1992, 737 (738)
255 *Foeste et al.*,a.a.O. (fn.177), § 81, Rn. 52.
256 Ebenda.
257 Ebenda.

た態度が存在する場合に生じる問題についても言及されるべきである[258]。このような諸事情では、誰が具体的な因果関係に寄与したのか、それが実際にこの人物に結果を帰属させるに十分であるのかを、常に正確に確かめるべきであろう[259]。この点において留意すべきなのは、回収決定における因果関係の問題である。例えば個別の取締役にとって、回収に賛成しても他の取締役は反対するため結局反対した、という場合である。このとき、その取締役については、義務を果たしていれば、結果が回避可能であったと言うことができない。そうすると、因果関係が認められないことになる。これが、問題となるすべての取締役について妥当する場合、結局、結果発生を誰にも帰属できなくなる。そこで、複数の直接原因をもたらしうる条件が、その全てを取り除いた場合に結果が回避できるとするならば、結果を回避ができるとして、因果関係が認められるという択一的因果関係のスキームを考慮することができる[260]。しかし、そのためには、回収決定前についても、過失の共同正犯を、回収決定後には故意の不真正不作為犯の共同正犯を認めざるを得ないのではないかという疑問が残る。

日本の不作為の因果関係の先例に関しては、最決平成元年 12 月 15 日刑集 43 巻 13 号 879 頁で「行為者の作為により十中八九結果を回避できたとするならば、刑法上の因果関係は認められる」と、不作為の因果関係の判断基準として合理的な疑いを超える程度に確実に結果発生を阻止できたか否かが分水嶺であることを示し、それと同一のスキームを用いた東京高判平成 21 年 2 月 2 日（刑集 66 巻 4 号 871 頁：三菱自動車車輪脱落事故控訴審判決）も「中国ＪＲバス事故の時点でDハブをリコールしてＦハブを装備しておけば、輪切り破損事故の発生はほぼ回避できたといえる」と判示する。しかし、事故原因を直接に惹起したとはいえない製造企業内部の品質管理部門責任者の結果帰属をこの因果的判断のみで認定するには過度な負担を課すことになろう。もっとも、前掲・最決平成 24 年 2 月 8 日では、「本件瀬谷事故は，Dハブを装備した車両についてリコール等の改善措置の実施のために必要な措置を採らなかった被告人両名の上記義務違反に基づく危険が現実化したものといえるから、両者の間に因果関係を認めることができる」と義務違反による危険の現実化という積極的判断にシフトしている。なお、同決定の田原裁判官反対意見における「一般に広く用いられている工業技術にかかる製品の瑕疵の有無及びその瑕疵に関する関係者の予見可能性の有無が基本的な論点となっている事件であり，その審理に当たっては科学技術的な観点からの十分な立証がなされるべき」という観点は重要である。というのも、同事件で問題となった事故原因については、長期使用の摩耗による可能性もありえたため、D ハブの摩耗によるものなのか強度不足に基づくものなのかが重要だからである。製造者に課せられる義務の観点からすると、品質保証をその目的とするリコール義務（道路運送車両法 63 条の 3）が課せられるのは強度不足に対応する義務のみであり、摩耗に対応する義務は製造者には

---

258 *Foeste et al.*, a.a.O. (fn.177), § 81, Rn. 53. 複数人の帰属については *Hassemer*, Produktverantwortung im modernen Strafrecht, S. 59 ff.

259 *Foeste et al.*, a.a.O. (fn.177), § 81. Rn. 54.

260 この点については、前嶋匠「企業・組織犯罪における合議決定と帰属関係(二・完)：因果関係と共同正犯・共同教唆」関大法学 54 巻 5 号（2005 年）166 頁以下を参照。

課されない（むしろ、利用者に課される点検・整備義務である）[261]。その意味では、製造者に課される義務に懈怠したとしても、直ちに刑事責任が帰属されるという状況にはならない。

まとめると、法律上の義務は刑法上の作為義務を導く根拠であるが、それだけでは十分でなく、刑法上の問題を検討するには、当該義務違反と結果発生との因果関係の検討も必要である。この観点は、製造者に対して、義務違反自体で直ちに刑事責任を負責させないという意味で、過度な負担を強いないという当初のテーゼとも調和する。そして、この因果関係の証明に際しては、第 1 章でも述べたような**説明可能な AI(explainaple AI)**の作成に努める義務を製造者に課すことが重要であると思われる。なぜなら、因果関係の証明において「因果関係のブラックボックス」の理論により、その因果経過の具体的内容が明確にならなかったとしても、刑事責任（過失責任）が製造者に負責される余地を残すので、結果的にこのことは製造者にとって負担になる可能性があり、AI 製品の開発・普及を阻むものとなりうるからである。

（7）帰属阻却要素

客観的帰属論は結果帰属の制限をねらいとするが、この帰属連関は、刑法上の製造物責任の領域で、以下の諸事情によって遮断されうる場合がある。例えば、利用者が、製造者の指示に注意を払わない、もしくは AI 製品のリコールがあるにもかかわらず、さらに利用し、それによって損害を被った場合である。これら事例では、被害者の自己答責性が考慮される[262]。しかし、その損害が利用者においてではなく、第三者において発生した場合、製品の危険は製造者の態度によってではなく、利用者の態度によってもたらされたものであり、それゆえ、製造者の刑事製造物責任は問題とならない[263]。同様のことは、製品の濫用、すなわち利用者が製品を故意に製造者の目的に反して使用した場合にも妥当する[264]が、このことに関しては第 4 目で検討を行う。

第 2 目　技術サービスプロバイダ

自動運転車の場合、当該システムの運用には、高精度の道路地図や車外のデータネットワークへのアクセス利用なくして有意義なものはなりえない[265]。そのため、上述した自動車が

---

261 この着想は注意規範保護目的論と親和的である。安達光治「危険の現実化論について」井田良ら編『浅田和茂先生古稀祝賀論文集（上巻)』（成文堂、2016 年）63 頁以下参照。また、松宮孝明「判批」立命館法学 343 号（2012 年）616 頁は、「本決定は、従来から『客観的帰属論』のひとつの具体化として主張されてきた、このような『保護目的』ないし『保護範囲』の考え方を、実質的に採用した初の最高裁判例であるといえる」という。

262 松宮・前掲（注 135）46 頁、130 頁。

263 *Foeste et al.*,a.a.O. (fn.177), § 81, Rn. 60

264 *Foeste et al.*,a.a.O. (fn.177), § 81, Rn. 63.

265 Vgl. *Sander/Hollering*, Strafrechtliche Verantwortlichkeit im Zusammenhang mit automatisiertem Fahren, NStZ 2017, S.199.

関与する交通事故におけるこれまでの刑事責任の担い手に加えて、それに対応する技術サービスプロバイダとその従業員にも焦点を当てるべきである。なぜなら、完全に自動化された車両の運転中の事故は、特に非常に正確であるはずの道路地図が不正確な表示をしたことや、環境の状態、交通インフラの事情、他の車両に関する情報など、車両に情報を提供することになっている情報・データネットワーク[266]の遮断が原因となりうる。従来の形式で制御された車両において、例えば、データ接続を使用してナビゲーション[267]や娯楽用電子機器類の領域で操作を容易にするなどの目的で、外部データにアクセスするのとは異なり、このようなデータとその受信は、多くの場合、車両の安全な移動にとって重要な意味を有する[268]。例えば、デジタル地図における現在のルート経過に関する情報が誤っていたり、データネットワークの障害により交通障害に関する情報が得られなかったりすることが交通事故の原因となることがある。このようなエラーが予測可能かつ回避可能な形で人の死傷に繋がった場合、地図表示の不備や重大なネットワークの障害に責任を負う技術サービスプロバイダの従業員は、このエラーの可能性を知ることができる限りで事故の結果の予見可能性を充足するといえよう。もっとも、回避可能であるか否か、さらには当該エラーに対応しなかったという不作為が結果発生の危険を現実化させたと言えなければ業務上過失致死傷罪を構成すべきではない。ただし、業務上過失致死の過失の内容である注意義務（作為義務）における、どの範囲まで技術サービスプロバイダが上記エラーに対応しなければならないのかについてはまだ開かれたままである。そのため、注意義務の内容をみだりに拡張させないようにするために、上記エラーを監視したり、場合によっては防止・修正したりすることを法律で義務付けることが必要であると考える。

第 3 目　国・地方公共団体（許可責任者）

危険な製品に対して、その流通権限の管轄たる国・地方公共団体における刑事製造物責任が問題となることがある。この具体的な事案としては、薬害エイズ（厚労省ルート）事件（最決平成 20 年 3 月 3 日刑集 62 巻 4 号 567 頁）が挙げられる。

その判示では、本件非加熱製剤の「危険性にかんがみれば，本来その販売，使用が中止され，又は，少なくとも，医療上やむを得ない場合以外は，使用が控えられるべきものであるにもかかわらず，国が明確な方針を示さなければ，引き続き，安易な，あるいはこれに乗じた販売や使用が行われるおそれがあり，それまでの経緯に照らしても，その取扱いを製薬会社等にゆだねれば，そのおそれが現実化する具体的な危険が存在していたことなどが認められる」。「このような状況の下では，薬品による危害発生を防止するため，薬事法６９条の２の緊急命令など，厚生大臣が薬事法上付与された各種の強制的な監督権限を行使するこ

266 *Sander/Hollering*, a.a.O. (fn.265), S.200.

267「ナビゲーターは補助的に使うことは許されるが、頼りにしてはならない」とライン川で従来のナビゲーションシステムがドライバーを誤誘導した実例を述べた *Joerden,* Strafrechtliche Perspektiven der Robotik, in: *Hilgendorf/Günther* (Hrsg.), Robotik und Gesetzgebung, Nomos 2013, S. 195 参照。

268 *Sander/Hollering*, a.a.O. (fn.265), S.200.

とが許容される前提となるべき重大な危険の存在が認められ，薬務行政上，その防止のために必要かつ十分な措置を採るべき具体的義務が生じたといえるのみならず，刑事法上も，本件非加熱製剤の製造，使用や安全確保に係る薬務行政を担当する者には，社会生活上，薬品による危害発生の防止の業務に従事する者としての注意義務が生じたものというべきである」（下線は筆者）と、必ずしも法律上の強制監督措置だけではなく，任意の措置を促すことで防止の目的を達成することが合理的に期待できるときは，そのような措置も防止措置に含まれるべきとした。

許可責任者の刑事責任として、最初に製品流通を可能にするために認定を行った者については、その認定によって新たな危険が生じる際にこれを防止する法律上の義務が存在し、これを懈怠した限りで可罰的となりうる。このような責任者に対する刑法上の過失責任は、例えば製造者に代わって作成された試験報告書など、認定手続の枠組で提出された書類を批判的に評価した際に、自動化技術、データ接続、使用されている地図材料などが記載されている方法では欠陥があること、ないしは、許可されるべきシステムの安全性に関する確かな検査が可能でなかったことを、技術的な専門知識を背景に知ることができたにもかかわらず、これを与えたという事例に認められよう。このようにして許可された製品の欠陥により、人間が死傷した場合、その許可責任者は過失犯における予見可能性を充足することになる[269]。もっとも、技術サービスプロバイダの場合と同様に、義務違反に基づく危険を結果発生に現実化させたと言えなければ業務上過失致死傷罪を構成すべきではない。

第 4 目　利用者

自立学習をする AI システムを直接利活用する利用者や所有者が、当該 AI 製品から他者の人命や身体を侵害した場合における刑法上の評価について分析する。具体的には、当該 AI 製品の利用により他者の人命や身体を侵害した場合、どこまで製造者側の指示・用法を信頼できるかが問題となる[270]。

（1）利用者の行為

利用者側の行為で問題となるのは、システムのアクティブ化が積極的作為としてみなされるか否かである。しかし、不作為は結果発生の不阻止において見られるものであるかもしれない。利用者がシステムのアクティブ化により再度積極的となり、積極的となったことが刑法上の結果に繋がるならば、積極的な行為が見て取られるだろう。その一例は、結果発生の直前の運転経過への介入や、利用者の入力を必要とする（部分的に）自律的なシステムの

269 自動運転車（レベル 2）の事例に限るが、*Sander/Hollering,* a.a.O. (fn.265), S.199 を参照。

270 この問題意識は *Valerius*, Strafrechtliche Grenzen lernender KünstlicherIntelligenz, GA 3/2022, S.121 にもある。もっとも、AI 製品の「利用者」の責任を具体的に論じるドイツの先行研究は少なく、分量を割いて検討しているのは *Günther*, a.a.O. (fn.176)（ロボット製品一般）や *Sander/Hollering*, a.a.O. (fn.265)（自動運転車）に限られる。

制御ないしは指揮である。しかしながら、AI 製品において通常想定されることが介在しない場合は、むしろ不作為が想定されうる。たとえシステムのアクティブ化をもって積極的作為を認めようとしても、発生結果に対する故意が欠如するため失敗してしまうだろう[271]。

（2）不作為における保障人的地位

利用者の不作為によって刑法上の結果が発生したという事例を基調にすると、AI 製品の利用形態を考慮すると、保障人的地位を保護的地位と監視的地位の間で分類する機能的二分説が役割を果たすものと思われる。先述のように、保護的地位とは、行為者が法益のために一定の義務を有し、外的な危険から被害者ないし被害法益を保護しなければならない地位であるため正犯を基礎づける一方、監視的地位とは、行為者が、危険源の拡大を監視し、そこから生ずるあらゆる侵害を防止する地位であるための共犯を基礎づける。保障人的地位は、保護的保障人に関しては常に被害者と行為者との関係について、監視的保障人の場合においては行為者と危険源との関係について判断されるが、どのような事例でこれらに該当するかを以下素描する。

① 保護的保障

保護的保障人の任務とは、保護すべき人格をめぐる外界からの危険から身を守る「防御盾」を形成することであるため、AI を搭載した介護用ロボットの利用の場合では、介護用ロボットの運用者に保障人的地位が生じうる[272]。注目すべきは、契約ないしは引受による保障人的地位である。ここには、行為者が他者のために保護義務を引き受けているといえる。例えば AI ロボットを社会福祉施設やサナトリウム、病院において利用することであり、人の移動やロボットを通じた治療における看護スタッフや医師は、患者にとっての保障人として考慮される。

② 監視的保障

監視的保障人は、危険が広がらないように、危険源の周りに「防御盾」を立てるべきとされる。その際、ロボット領域に関連するのは、特に危険源としての物についての責任と、過去の作為による保障人的地位であると考えられる。物的危険に関しては、利用者は製造者側が提供する指示に従って AI 製品が他人に危険を生じさせないように維持・監視する必要が

---

271 行為と結果の間に介在する故意という限りで,考えられうる後発的故意[*dolus subsequens*]の事例につき,後発的故意についても含めて一般的には *Jerouschek/Kölbel*, Zur Bedeutung des so genannten Koinzidenzprinzips im Strafrecht, JuS. 2001, 417 ff を参照。もっとも、後発的故意での処理の可能性も考慮されるが、結果発生の直前の瞬間に故意を認めるのは疑問が残る。むろん、AI システムの利用において作為犯構成は困難であり、利用者はそのシステムに身を委ねている以上、不作為の問題となりうる可能性が高い。

272 ただし、想定事例においては Y の X に対する関係に限ることに注意されたい。

あろう[273]。その際、危険性が社会的に相当であるか、ないしは義務に従ったものであるか、それとも、義務違反に基づいているか決定的ではない[274]。自動車の所有者や占有者、動物の飼育者の場合[275]、特定の危険を防止するための保障人的地位に基づき、AI 製品の利用者の場合も特殊な保障人的地位が想定される。さらに、義務違反的に危険を生じさせるような先行行為も保障人的地位になることがある[276]。なぜなら、その人がこのような切迫した危険を生じさせた場合、第三者の法益への危険を回避する義務を負うからである。利用者による先行行為の事例に該当するのは本来の用途に反して機器を使用したり、製造者の指示に従わず当該製品を濫用したりした場合に限定される。

（3）利用者に課せられる義務

注意義務は、法規範もしくは規則などの準則から生じることもありうるが[277]、AI・ロボットの利用にとってそのような準則は現在のところほとんど存在しない。このような準則によって具体化されなければ、他人への侵害を阻止するために、利用者にどのような要件を課さなければならないかを確認するための個別事例における準則を作成しなければならないことになる[278]。その際、構造方式とロボットの使用目的や適用領域が基準の決定に重大な影響を与えうる。構造方式、構成、製造により危険が予想される場合は、注意義務の程度が高められるし、システムがどこで投入されるかもその注意義務の具体化の決め手となりうる。これらは製造者側が想定する指示・規定に大きくよるところであり、これに違反して AI 製品を使用すると、より高い危険性が考慮される。さらに、その AI 製品の挙動については別途考慮しなければならない。もし当該 AI 製品が同様の状況で疑わしい挙動をしたことがあれば、利用者は製造者にこの挙動について報告するなどの措置を講じさせるに留めるべきであろう。というのも、利用者に対してこの場合に格別高い注意を求めるのは当該製品の技術的側面に照らせば過大な負担を課すことになってしまうからである。ただし、製品事故に最も近接する主体として、利用者は少なくとも製造者側に対しさらなる結果発生を防止するためこのような情報を適時に伝える義務を課すべきと考える。さらに、当該注意義務の確定においては、利用者の特性を考慮することも重要であろう。例えば、行為者がある社会的な人的集団に分類される場合、ここには、それと同等の人物が AI・ロボットとどのような

---

[273] *Joerden,* a.a.O. (fn.267), S. 209 では、観察は恒常的で適切なものでなければならないと言う。しかし、このことはすでに「自立」システムの意味が失われないのだろうか否かという疑問が提起される。

[274] *Wessels/Beulke,* Strafrecht Allgemeiner Teil, 52. Aufl., C.F. Müller 2022, Rn. 723.

[275] *Fischer*, a.a.O. (fn.181), § 13. Rn. 64. m.w.N

[276] *Wessels/Beulke*, a.a.O. (fn.274), Rn. 725. m.w.N.

[277] 下位の準則から義務を生じることはあるが、それが直ちに刑法上の注意義務違反（過失）に繋がるわけではない。もっとも、客観的帰属論における「許されない」危険創出における、「許されない」には下位規範からも導かれるものである。また、規範の保護目的についても考慮せねばならない。注意規範の射程内で起こったことについては帰属させるべきであるが、射程外のものについては帰属させるべきではない。それゆえ、第 1 目でも言及した客観的帰属論は注意規範の精緻化を導くものであるといえる。

[278] Vgl. OLG Hamm, Beschluss vom 5.1.1996, Az.: 2 Ss 1035/95, NJW 1996, 1295.

経験をしてきたか、またはそのような者がどのような知識を有しているかも客観的に決定されるべきである。確かに、これらの点は具体的な事例形式に応じて異なる重み付けをすることができるが、これらの基準を考慮して初めて、利用者に課せられる義務を明確にできる。ただし、この義務が決して最終的なものではないということに注意しなければならない[279]。なぜならば、義務違反は過失を認定することの十分条件でしかないからである。

（4）因果関係と客観的帰属

想定事例における利用者にとっては、課せられた作為義務の履行により危険を結果発生に現実化させなかったといえることが因果関係の結節点となる。そこで問題となるのは、法的基準に加えて、因果関係の具体的な証明である。このことについて、因果関係のブラックボックスの導入により利用者の行為と結果の間に因果関係が認められることもある[280]。

これら上記のスキームに照らし、冒頭の想定事例を再度検討すると以下のようになる。まず、①事例において、X は当該介護用ロボットの直接利用者であり A 以外の主体に対する監視的保障が認められうるが、当該製品の一方的受益者であるので結果回避可能性を欠く可能性が高く過失を構成するのは難しい。それに対して Y は、製造者 Z の指示がある限りで、直接の利用者ではないが X と A に対する監視的保障義務が認められ、その義務の不履行による危険がC の死亡結果に現実化したと認められる限りで過失致死罪が肯定されうる。②事例におけるWはその管理権限に照らして、Bに対する保護的保障義務が認められうる。そして、その義務の不履行による危険が B の死亡という結果に現実化したといえる限りで業務上過失致死罪を構成するといえよう。

第 5 目　所有者

最後にこの項では AI 製品の所有者と利用者が異なり、利用者がその使用中に死傷したという事例を想定した場合の所有者の刑事責任について検討する。

具体的には以下の事例を想定する[281]。オートパイロットを搭載した F 保有の自動車に乗車していた G は、その作動中、車内に搭載されていたモニターで映画を視聴していた。同車が高規格道路を法定速度で走行していた際、右方から来たトラックが道路を横切ったとき、G もオートパイロットも自動車のブレーキをかけなかった。そのため、同車はトラックに衝突し、トレーラーの下に滑り込み、フロントガラスと屋根が押しつぶされ G が死亡した。自動操縦カメラもレーダーシステムも、衝突の危険を認識していなかった。事故後の調査では、トラックを交通標識と誤認識したか、あるいはその白い側壁と背後の昼間の空を区別できなかったという。

この事例においては、まず車両の所有者の刑事責任が問われうる。ここでは車両所有者は

279 *Günther*, a.a.O. (fn.176), S.213.
280 たとえば、前掲・最決平成 12 年 12 月 20 日が挙げられる。
281 *Sander/Hollering*, a.a.O. (fn.265), S.196.

利用者に対する過失致死罪の適用が問題となるが、車両に対する実際の処分権を有する者としての車両所有者の可罰性のポイントは、やはり所有者に課せられる注意義務と、その義務の不履行による危険に現実化したかということにある。まずは、考慮される注意義務を列挙すると、以下の義務が想定される。まず、運転手への不十分な指示（指示の欠如）である。所有者は、例えば自動運転システムの使用方法を事前に、十分にドライバーに教えていなかった場合、義務違反的態度として非難されるべきであろう[282]。例えば、指導を怠った結果、ドライバーが予見可能かつ回避可能な形で自動車を適時にオーバーライドすることができなかった場合のように、指示の欠如が死傷結果に繋がったといえる場合は、過失致死傷罪の成立が考えられる。ドライバーが所有者の当該義務の不知を認識できる形で理解していなかった場合も同様である[283]。次に考慮されるのは、委譲された自動車の不十分な整備（整備不良）である。所有者側の注意義務違反は、完全自動運転システムを搭載した車両の整備が不十分であることも原因となりうる[284]。所有者には道路運送車両法 47 条により所有者についての整備義務が規定されているほか、同法 49 条 2 項では特定整備[285]に係る点検記載義務が課せられる。もし前述の事例で、F が規則に適合して整備された状態で車を運転手に預けたが、複雑な自動化技術の点で、特に重要なセンサー技術の分野では十分に考えられる欠陥が、F にとって予見可能かつ回避可能な態様で死亡事故につながった場合、F の注意義務違反を構成しうる。さらに、自動運転システムの機能安全性の欠如（機能不良）も考慮される。自動運転システムの設計やプログラムの欠陥による損害の場合、所有者は、自動運転システムが、それが単純な製造上の欠陥であるかによらず、その設計やソフトウェアによって安全ではない可能性があり、誤作動があれば結果的に非常に危険な運転状況を引き起こす可能性があることを認識した上で、ドライバーに車を委譲したという観点から、すでに過失として非難されることがあるか否かを検討すべきである。なぜなら、道路運送車両法 41 条によれば、車両の所有者は、車両が技術基準に適合していない場合には、車両の供用を許可してはならないからである。基準に適合しているということは、車両が道路運送車両法 40 条以下に対応していなければならず、また、安全に運転できなければならないということであるから、予期される設計やプログラミングに条件付けられた機能不良については、独立した義務違反が考慮される。

---

282 *Ebenda*.
283 *Ebenda*.
284 Vgl. *Sander/Hollering*, a.a.O. (fn.265), S.197.
285 道路車両運送法施行規則 3 条 8 号によると、自動運行装置に関連する特定整備として、
次に掲げるもの（以下「運行補助装置」という。）の取り外し、取付位置若しくは取付角度の変更又は機能の調整を行う自動車の整備又は改造（かじ取り装置又は制動装置の作動に影響を及ぼすおそれがあるものに限り、次号に掲げるものを除く。）
イ　自動車の運行時の状態及び前方の状況を検知するためのセンサー
ロ　イに規定するセンサーから送信された情報を処理するための電子計算機
ハ　イに規定するセンサーが取り付けられた自動車の車体前部又は窓ガラス
を規定する。

これら義務違反性が認められた上で、所有者 F が過失致死罪の罪責を負うとするには、F の義務違反という危険が G の死亡という結果発生に現実化させたといえるかが結節点となる。もっとも、G の自己答責的な自己危殆化という観点のもと、F が責任を負いえない場合もあることを考慮すべきである。車両を使用する際に、F が自動運転システムのありうる機能的欠陥から生じるリスクを認識して当該自動車の使用を禁止していたにもかかわらず G がこれを無視して当該自動車を使用した場合、車両所有者へのリスク実現の帰属可能性は欠如する。

第５項　許された危険による解決

たとえば自動運転車が普通自動車の有するリスク、すなわち普通自動車による交通事故件数よりも少なくなるというのであれば、自動運転車の有するリスクを社会が許容しうる可能性はある[286]。むろん、自動運転車に限らず、AI を搭載した他の製品についても同様に該当するものと思われる。

これにつき、過失においては、注意義務は許された危険という形相を通じて制限されうるとする Günther の見解を紹介する。彼は、許された危険の体系的位置づけにつき、「許された危険という形相は、注意義務において考慮されるべきであり、固有の正当化事由を表すものではない」[287]とする。そして、「その場合の出発点は、我々の高度に技術化されたリスク社会において一定の危殆化は完全に排除されうるものではない」ことである[288]ので、法秩序は、危険な態度においても、それに伴う社会的利益のために、典型的にそれに伴う危険の避けられない残存リスクを受容するものとする。AI については、製造者が経済的な側面を引き合いに出すことで、安全面で技術的に実現可能なことをすべて実現する必要がない。彼は AI や「ロボット工学の関連では、許された危険が特に重要な意義を持つ」[289]という。特に高齢者介護の領域で、ロボットの集中的な研究が行われていることを引き合いに、将来、様々な生活領域でロボットが登場し、ますます日常的なものとなると想定する。初期のロボットは、確かに性能は低いと言わざるを得ないが、すでに生産は進んでおり、私人でも入手可能となっているため、これらの製品の中には社会に著しい利益をもたらすものもあるから、ロボット工学がもたらす危険は、許された危険で顧慮されうるものもある[290]という。

例えば、製品が販売に値しないものにならないようになってしまう場合、製造者はその設計において、安全製造メカニズムを全て組み込む必要はないとする[291]。同様に、製造領域でも、製造者が十分な管理機関を創設し、それによって製造上の欠陥が排除されうる場合、彼は（そのために許容すべき）外れ値のために可罰的とはならないという[292]。しかしながら、

---

286 松宮・前掲（注 11）4 頁以下。

287 *Schönke/Schröder,* Strafgesetzbuch: StGB, 30. Aufl., 2019, Vorb. § 32 ff., Rn. 107b.

288 *Schönke/Schröder*, a.a.O. (fn.287), § 15, Rn. 144.

289 *Günther*, a.a.O. (fn.176), S.226.

290 *Günther*, a.a.O. (fn.176), S.226.

291 *Foeste et al.*,a.a.O. (fn.177), § 81 Rn 42 f.

292 *Günther,*a.a.O.(fn.176)*,* S.226.

許された危険を通じて、製品監視に関しては注意義務の制限を可能にすべきでないという[293]。その背景は、発生した危険に対処するために、その製造者全てがその可能性を用いるということにある[294]。許された危険を、生産や製造の分野での配慮を減らすことになる、経済的な考慮に基づく製造者のための容認として正確に理解するならば、これは製品監視義務者の側でより大きな注意を払わなければならない。このように、もはや経済的に支持できる生産の問題ではないため、許された危険によって弱められた基準を想定することはできない。むしろ、製造者は自ら有するすべての可能性を尽くさなければならない。特に、その危殆化はとりわけ製造者について条件付きで評価されるため、インテリジェントな機械の製造者の製品監視義務はさらに重いものとしてみなされるべきであるとする[295]。

確かに、利益衡量によって「許された危険」を認定することについては、功利主義的観点からは首肯しうるものであるものの、このような許された危険の適用については、どのような基準をもって「許されたもの」とするのかが問題となる。それは、第一に、規則遵守の有無、第二に注意遵守の有無に分類される[296]。第一の「規則」の射程としては、実定法上の規範だけでなく、主務官庁の認可の条件となった各種の保安設備や、就業規則も含まれるとする。第二の基準については刑法上であっても落ち度のない態度がとられたと判断されることがその内容であるという。その意味では、AI ロボット開発に関する諸規格やガイドラインの非具体性に鑑みると、不明確な基準で「許された」と言うには時期尚早であるように思われる。もっとも、仮に第一の要件を充足すれば、第二の要件たる、当該行為が「落ち度のない態度」と評価されるか否かについては、AI・ロボットによってもたらされる便益が、これによって生じる害や危険を上回り、社会的受容が得られる暁には肯定され、これによって初めて当該行為は「許された危険」となるだろう。

第５節　小括

自立学習をする AI 製品に起因する事故により人が死傷した場合、その AI 製品をめぐる製造者・利用者を中心とする各主体は刑法上の負責から完全に解放されることはない。しかし、その死傷結果はいかなる主体に帰属されるのかという問題は、製造者側については設計上の義務、製造上の義務、指示・警告上の義務、製品監視義務といった製造物責任上の義務を、技術サービスプロバイダにはネットワーク化された製品におけるエラーを監視、防止及び修正する義務を、国家や地方公共団体などの許可責任者には当該 AI 製品に対する一定の品質管理に関する法律上の義務を、利用者には製造者の指示に従う義務を、そして所有者には当該 AI 製品に関する製造者からの指示及び法律上の義務を明確にしないことには解決が困難であるということが明らかとなった。この義務内容として重要なのは AI 製品に対する

---

293 Vgl. auch LG Frankfurt, Urteil vom 25.5.1993, Az.: 5/26 KLs 65 Js 8793/84: それゆえ製造者は生命や身体への危険を警告しなければならない。

294 *Foeste et al.*, a.a.O. (fn.177), § 81. Rn. 45.

295 *Günther*, a.a.O. (fn.176), S.226.

296 西原春夫『交通事故と信頼の原則』（1973 年、成文堂）35 頁以下。

監視義務であり、これは利用者・製造者双方に関連する。ただし、この義務違反が直ちに刑法上の過失を構成するものとしてはならない。さらには、義務を履行していればほぼ確実に結果を回避できたとする仮定的因果関係のスキームを用いるべきでもない。むしろ、義務違反に基づく危険が死傷結果に現実化したといえて初めてこれを構成するものであるべきである。このスキームにより過失の処罰範囲のみだりな拡張を防ぐことが可能であるが、このことは、AI 製品の普及・利活用の観点から、製造者や利用者に過度な負担を課さないとする思想にも親和的である。ともあれ、今日において求められるのは、まずは製造者や利用者など AI 製品に関与する主体に課せられる義務―監視義務―を行政法規ないしはガイドライン等などにより明確化することにある[297]。

第 3 章　さらなる AI の利活用における刑法上の諸問題―財産侵害

第 1 節　問題の所在

第 2 章では日常的な AI 製品に由来する人命や身体に関する侵害事例における従来型の議論について検討したが、AI は生命や身体以外の保護法益、例えば財産そして経済領域における投資家の利益を侵害する可能性があることも第 1 章で示唆した。本章では、そのような経済犯罪、さらには AI 自体が行為客体となりうるコンピュータ犯罪における AI 製品に関する刑法上の問題について、その法律の制定背景ならびに構成要件の解釈を、具体的事例をあげながら検討を行う。

第 2 節　経済犯罪

相場操縦行為やインサイダー取引といった不公正取引やカルテル行為といった競争法違反の場合、過失犯処罰規定が存在しない。そのため、AI・アルゴリズムの利用の結果として不公正取引や競争法違反がなされた場合、果たしてその AI の利用者・管理者、開発者、販売者に対してどのようにして刑法上の責任を帰属されるのかについては別途検討を要する。その手法として、まずはその規定の概要をたどり、想定される事例について現行法の規定から検討を行い、現行法の解釈で足りるのか、それとも立法的解決が望ましいのかを検討する。

第 1 款　相場操縦行為

第 1 項　問題の所在

AI・アルゴリズム投資プログラムによって相場操縦・インサイダー取引が行われた場合に、その利用者・販売者・製造者はいかなる責任を負うのかという問題が提起されている。近時、例えば、複数の取引施設から最良価格を提示し、取引施設を検索し注文を執行する Smart Order Routing（SOR）というシステムの普及が指摘されているなど[298]、AI・アルゴリ

---

[297] 例えば中国の AI 開発規則である中华人民共和国科学技术部「新一代人工智能伦理规范」（2020 年）17 条では、製造者や利用者における義務が明文化されている。

[298] 金融審議会市場制度ワーキング・グループ「最良執行のあり方等に関するタスクフォース 報告書」

ズムを利用した高速取引、大量取引が広く行われるようになっている。このような高速取引行為については、市場に流動性が供給されているとの指摘や、流動性が厚くなることでスプレッド（価格差）が縮まり一般投資家にもその恩恵が及んでいるとされる。
しかし、アルゴリズムを用いた相場操縦等の不公正取引の事案等が報告されているなど、市場の公正性に影響を与えるおそれも指摘される[299]。例えば、北越紀州製紙の株式で、実際には成立させるつもりのない売買注文を「見せ玉」として発注し、コンピュータが自動的に発注を繰り返すアルゴリズム取引により株価を1〜2円単位で上下させ、不正に利益をあげたことにより、相場操縦として、デイトレーダーに対し57万円の課徴金納付を命じた事例[300]が存在する。このような懸念に対応する形で、金融商品取引法（以下「金商法」とする）2017年改正では、高速取引業者に関する諸規定が創設された（金商法2条41号、2条42号、金商法60条の55など）。

この AI・アルゴリズムの投資判断に関する学習過程のうち、実データを使用して「もし取引していたらどうなっていたか」を分析するバックテスト方式を利用していた場合、自身の取引が市場に与える影響を考慮できないゆえに、バックテスト方式である限りは AI が相場操縦をする心配はない。しかし、仮想市場をコンピュータ上に作成し分析する人工市場を用いたシミュレーションでは、「自らの取引が市場価格に与える影響を継続的に学習・分析する」ことは可能ではないかという仮定のもと、仮想市場モデルでは、相場操縦にほかならない取引（仮想売買型、見せ玉、現実取引型を問わず）を最適な取引として導出したという結果が報告されている[301]。このことから、AI・アルゴリズムによる投資判断において、相場操縦取引を行う可能性を利用者・開発者は認識することが可能となる。そうすると、金商法159条および157条の要件該当性に影響を及ぼすものと考えられる。これを踏まえて、上記論文や先行研究[302]では金融商品取引における AI・アルゴリズムの開発者にとって、相場操縦をするようなプログラムをしない義務を新たに課す必要性への示唆がある[303]。本項では、現行法下での解釈とその妥当性、そして必要があれば立法的解決が必要であるか否かを検討する。

第2項 相場操縦規制の概要

相場操縦規制の趣旨は以下の通りである。すなわち、投資者は金融商品市場における価格が公正にして自然な状態における需要と供給によって形成されたものであると信用するた

（2021年）2頁。

299 金融審議会市場制度ワーキング・グループ・前掲（注298）10頁。

300 日本経済新聞「金融庁、アルゴリズム取引悪用の相場操縦で課徴金命令」（2011年2月16日）

301 水田孝信「人工知能は相場操縦という不正な取引を勝手に行うか？―遺伝的アルゴリズムが人工市場シミュレーションで学習する場合―」第34回人工知能学会全国大会論文集（2020年）1頁。

302 アルゴリズム・AI の利用を巡る法律問題研究会・前掲（注4）20頁以下。以下、この文献を「先行研究」ということがある。

303 水田・前掲（注301）4頁。

め、人為的な操作によって高騰あるいは下落させられた相場を公正な相場だと誤認させることは、投資家の期待を裏切るものとなる。これにより、相場操縦に関する規制を設け、正常かつ自然の需給関係による市場を保証しようとするのが当規定の趣旨である。アメリカ1934 年証券取引所法 9 条を参考に、(旧) 証券取引法 125 条として規定されたものである[304]。近時、徐々に摘発例が増えてきており、デイトレーダーがネット取引による「見せ玉」等の手法を用いて相場操縦を行う事例が報告されている[305]。

主な禁止行為類型は以下のように分類される。第一に「取引が繁盛に行われていると他人に誤解させる目的やその他のこれらの取引の状況に関し他人に誤解を生じさせる目的」をもった取引として、仮装売買[306]（金商法 159 条 1 項 1~3 号)、馴合売買[307]（金商法 159 条 1 項 4~8 号)、仮装売買・馴合売買の委託または受託（金商法 159 条 1 項 9 号)。第二に、「取引を誘引する目的」をもった現実売買[308]（金商法 159 条 2 項 1 号）第三に、「相場をくぎ付けし、固定し、又は安定させる目的」をもってする安定操作（金商法 159 条 3 項）である。その他にも、「空売り」(金商法 162 条 1 項 1 号）や、約定させる意思がないにもかかわらず市場に注文を出し、取引が成立しそうになると注文を取り消す「見せ玉」もこの相場操縦に該当する。とりわけ、この見せ玉には、未執行の注文動向をインターネット上で知ることが可能となったという背景がある。なお、どの類型に該当するかについて、投資家による見せ玉は「委託」に該当するとされ、金商法 159 条 2 項 1 号で把握される。(この条文における「申込み」が見せ玉類型にあたる)。ただし、変動取引の連続性も要件に含まれるので、1 回限りの見せ玉では金商法 159 条 2 項 1 号に該当しないし、見せ玉の違法性は虚偽の注文を作出して一般投資家に誤解を与えることにつき、この行為は金商法 159 条 1 項の仮装売買・馴合売買型と同一視すべきであると批判される[309]。

その法効果としては、10 年以下の懲役もしくは 1000 万円以下の罰金またはその併科（金商法 197 条 1 項 5 号）さらに課徴金（金商法 174 条〜174 条の 3）がある。さらに、刑事罰

---

304 この歴史的展開については,張小寧「証券犯罪の総合的研究（2）—実効的規制のための基礎的考察—」立命館法学 343 号（2012 年）59 頁以下が詳しい。

305 神田秀樹・黒沼悦郎・松尾直彦編『金融商品取引法コンメンタール 4 不公正取引規制・課徴金・罰則』(商事法務、2011 年）23 頁（藤田)。

306 権利の移転を目的としない仮想の株式売買を指す。例えば,ある投資家 X が A 証券会社に対し S 社株を 1 株 1 万円で買い注文を出す一方で、同じ S 社株を B 証券会社に対し 1 株 1 万円で売り注文を出す。そうすると、S 社株の相場を 1 株 1 万円に形成することができる。なお、この行為主体については、条文の規定が「何人」とされているところ、法人も含まれることに留意しなければならない。斉藤豊治・浅田和茂・松宮孝明・髙山佳奈子編『新経済刑法入門 (第 3 版)』(成文堂、2020 年）200 頁（平山）参照。

307 馴合売買とは注 306 の事例において、複数人が関与する場合である。具体的には、売り注文と買い注文をする主体が別の主体であり、その中で通謀がなされ、かつ同時期に同価格で売り注文と買い注文をことが必要である。斉藤ほか・前掲（注 306）200 頁（平山）参照。

308 例えば、X は他の投資家を A 社株の売買取引に誘引する目的で同社株式の大量の買い注文を出すことにより A 社株式の株価を上昇させた事例が挙げられる。

309 黒沼悦郎『金融商品取引法[第 2 版]』(有斐閣、2020 年）502 頁。

につき両罰規定がある（金商法 207 条）。具体的には、法人の代表者又は法人若しくは人の代理人、使用人その他の従業者が、その法人又は人の業務又は財産に関し、その行為者を罰するほか、その法人に対して 7 億円以下の罰金刑が科されるというものである。

第 3 項 AI・アルゴリズムを用いた取引と相場操縦規制

AI・アルゴリズムを利用した投資判断において、どのような禁止行為が問題となるのか。まずは、高頻度取引(HFT)と AI・アルゴリズム取引の定義から行う。「高頻度取引（High Frequency Trading）」[310]とは、コンピュータを駆使した超高速の金融取引を指す。過去の価格の動きを統計的に分析し、1 秒間に数千回もの高頻度で売買の注文を繰り返す。わずかな値幅、瞬時の動きをとらえて資金を回転させることで利益を積み上げる。この分析プロセスについて、予めプログラムされた通りの分析をするものもあれば、AI を組み込み、学習を重ねた上で分析を行うものも存在する。主に後者のものを「AI・アルゴリズム」取引として定義する。

金商法 159 条 1 項所定の禁止行為を、仮装売買等をするように設計した AI・アルゴリズムを介して投資家が意図して行った場合は要件解釈上の問題が生じることはない。しかし、AI・アルゴリズムの「学習」によって新たに投資判断プロセスが構築され、利用者（投資家）が知らないまま、それが結果として 159 条 1 項・2 項所定の禁止行為がなされた場合についてはどのようになるか。その AI・アルゴリズムの学習過程が人間によって理解できないものである、すなわちブラックボックス化している際、AI・アルゴリズムは、自らで一定の相場変動をもたらす取引を行うことで利益を得るような戦略を採用したように~~は~~見える。それでは、AI・アルゴリズム自身が取引主体となるように見えるが、AI・アルゴリズムは「人格」ではないため、この禁止行為規定をそのまま適用することはできない。ゆえに、AI・アルゴリズムの背後に存在する利用者について、この規定が適用されるか否かを類型ごとに検討する必要がある。

第 1 目 仮装売買・馴合売買類型

この AI・アルゴリズムを利用した主体にとって、取引が繁盛に行われていると他人に誤解させる目的やその他のこれらの取引の状況に関し他人に誤解を生じさせる目的を認定することはできるか。それにはまず、当該規定における「目的」の判例上の解釈にさかのぼって検討する必要がある。

（1）目的規定の解釈

判例の見解は以下のようなものである。すなわち、上記の「目的」とは、「取引が頻繁かつ広範に行われているとの外観を呈する等、取引の出来高、売買の回数、価格等の変動ならび

[310] 日本経済新聞「高速取引(HFT)とは　データ基に 1 秒で数千回の売買注文」（2019 年 10 月 20 日）

に参加者等の状況に関し、投資者に自然の需給関係によってそのような取引の状況になっているものと誤解されるものであることの認識」[311]とされる。また、「出来高に関し他人に誤解を生じさせる目的も、上記『取引が繁盛に行われていると誤解させる等これらの取引の状況に関し他人に誤解を生じさせる目的』に当たり、**特定の銘柄についての価格操作ないし相場操縦の目的を伴わない場合でも、本罪**（159条3号及び8号―筆者注）**は成立する**と解すべきである」[312]（太字・傍線筆者）というように、特定の銘柄についての相場操縦をする目的がなくとも、単に何かしらの有価証券等の相場操縦をする目的があれば足りるものとしている。

また、取引の状況に関する誤解を生じさせる目的があれば足り、後述する現実売買による相場操縦の場合とは異なり、有価証券売買等を誘引する目的、最終的に相場を変動させる目的があったか否かは必要としない。さらに、当該目的は「仮装売買をすること自体が、特段の事情のない限り、取引の状況に関し他人に誤解を生じさせる目的を強く推認させるものである。この目的があるというためには、被告人が、自身が行おうとしている取引を行えば第三者がその取引状況に関し実需に基づくものであると誤解する可能性があることを認識した上で、当該取引を行ったことが認められれば足りるというべき」[313]というように推認的に目的要件を認定することもありうる。

（2）AI・アルゴリズム投資と仮装売買型相場操縦規制の検討

AI・アルゴリズムの利用者にとって、その投資判断の評価過程がブラックボックス化する限り、投資者に実際に仮装売買が行われたからといって投資者に遡って相場操縦罪の罪責を帰属させるのは投資者に予測不可能なリスクを負わせることとなり、AI・アルゴリズム投資の便益を損なうおそれがある。むろん、AI・アルゴリズムの予測できない挙動により、仮装売買を有効な投資判断としてこれを行う可能性は考慮されうるし、先述のような非常に制限された条件下ではあるが、AI・アルゴリズムが相場操縦を最適な取引として検出したというエビデンスも確かに存在する。しかし、そのような限られた可能性のみで取引が繁盛に行われていると他人に誤解させる目的やその他のこれらの取引の状況に関し他人に誤解を生じさせる目的を認定するのは投資者にとって酷であろう。このことは、当該AI・アルゴリズムの開発者である製造者にも妥当する。

結論として、現行法の規制では、AI・アルゴリズムの利用者や製造者に対して仮装売買型・馴合売買型の相場操縦規制に基づくエンフォースメントを行使するのは困難ということになる。

---

311 大阪地判平成20年10月31日 裁判所ホームページ。

312 最決平成19年7月12日 刑集61巻5号456頁。

313 大阪地判平成18年7月19日 裁判所ホームページ。

第 2 目　現実取引型相場操縦

AI・アルゴリズムの「学習」によって新たに投資判断プロセスが構築され、利用者（投資家）が知らないまま、それが結果として 159 条 2 項所定の禁止行為がなされた場合はどのようになるか。この点について、159 条 1 項の場合と要件が異なるため、ここでもまずはその法的性質並びに要件についての検討を行う。

（1）法的性質とその要件

市場における株式の価格は、大量の売り注文・買い注文を継続的に出すことによっても、人為的に形成することができる。しかし、たとえ大量の株式の売買であり、相場を変動させるものであっても、投資者の必要に応じて正当に行われている行為を規制することはできない。そこで問題は、いかにして適法な取引行為と違法な取引行為を区別するかにある。金商法 159 条 2 項 1 号は、主観的要件としての「取引を誘引する目的」（誘引目的）をもって、客観的要件としての「相場を変動させるべき」取引（変動取引）を行うことを禁止の対象とすることにより、適法な売買行為と区別しようとする。

まず、誘因目的（主観的要件）とは、「有価証券の相場を変動させるべき一連の売買取引等のすべてを違法とするものではなく、このうち『有価証券市場における有価証券の売買取引を誘引する目的』、すなわち、人為的な操作を加えて相場を変動させるにもかかわらず、投資者にその相場が自然の需給関係により形成されたものであると誤認させて有価証券市場における有価証券の売買取引に誘い込む目的」[314]と解される。次に相場を変動させるべき取引（客観的要件）とは、「相場を変動させる可能性のある売買取引等」[315]と解される。より具体的には、「『相場を変動させるべき取引』とは、同号が売買取引のほかその委託、受託をも併せて禁止していることに徴し、市場価格を変動させる可能性のある取引を広く指称する」と解される[316]。

この最高裁決定により、誘因目的が違法／適法の評価を画するものとしたものの、その目的の有無は客観的な事情から推認せざるを得ないことが多い。例えば、東京地判平成 5 年 5 月 19 日判タ 817 号 221 頁（藤田観光株株価操作事件）では、「この目的の存否は、もちろん当事者の供述からそれが明らかにできることはあるが、そうした供述によることなく、取引の動機、売買取引の態様、売買取引に付随した前後の事情等から推測して判断することは十分可能であり、その際には、売買取引の態様が経済的合理性をもったものかどうかが、人為的に相場を操作しようとの目的を窺わせるものとして、重要な意味を持つといえる」として

---

314 最決平成 6 年 7 月 20 日刑集 48 巻 5 号 201 頁[協同飼料事件上告審決定]。原審（東京地判昭和 59 年 7 月 31 日判時 1138 号 33 頁）「『売買取引を誘引する目的』とは、市場の実勢や売買取引の状況に関する第三者の判断を誤らせてこれらの者を市場における売買取引に誘い込む目的、すなわち、本来自由公開市場における需給関係ないし自由競争原理によって形成されるべき相場を人為的に変動させようとの意図のもとで善良な投資家を市場における売買取引に参加させる目的をい」うという見解を維持している。

315 最決平成 6 年 7 月 20 日刑集 48 巻 5 号 201 頁[協同飼料事件上告審決定]。

316 東京地判昭和 59 年 7 月 31 日判時 1138 号 33 頁。

いる。誘因目的でなされた「変動取引」と認定される例[317]としては、合理的な理由なく、終値付近に大量の買付けを行う（終値関与）、1円ごとに指値を高くする注文を大量に出す（買い上がり）、（空売りで）安値の売り注文を大量に出したりする行為（売り崩し）が挙げられる。

（2）AI・アルゴリズム投資と現実取引型相場操縦規制

AI・アルゴリズムを利用した取引において現実取引型相場操縦規制はどのように適用されるか。利用者が誘因目的をもって変動取引を行うような AI・アルゴリズムを構築した場合は、当然に 159 条 2 項型の相場操縦規制違反となろう。かりに、当該利用者が、その AI・アルゴリズムが他の投資者に対して、その相場が自然の需給関係により形成されたものであると誤認させてしまう取引を行うものであることを認識していたにもかかわらず、取引をなお継続させていた場合でも利用者の誘因目的を認定することは可能であろう。

しかし、AI・アルゴリズムを利用する際、取引が相場に与える影響を継続的に学習・分析し当該分析に基づいて取引を行うことができるため、ともすれば利用者、たとえば法人の代表者等の知らないところで、AI・アルゴリズムは自立的判断により、一定の相場変動をもたらす取引を行うことで利益を得るような戦略を採用することも考えられる。この場合「変動取引」は認定しうるとしても、利用者の「誘因目的」を認定することは困難である。このことは、AI・アルゴリズムの開発者たる製造者にも該当する。そうすると、159 条 2 項の規制の適用は難しいだろう。そのため、このような相場操縦を未然に防ぐために、金商法における他の規制、例えば業者規制の適用可否についても検討しておく必要がある。

第 3 目　業者規制（金融商品取引業者・高速取引業者）

金融商品取引業者には金商法 40 条 2 項により以下の業者規制が課せられる。すなわち、金融商品取引業者等は、業務の運営の状況が「業務に関して取得した顧客に関する情報の適正な取扱いを確保するための措置を講じていないと認められる状況、その他業務の運営の状況が公益に反し、又は投資者の保護に支障を生ずるおそれがあるものとして内閣府令で定める状況にあること」に該当することのないように、その業務を行わなければならない、と。そこでいわれる「状況」とは、金融商品取引業等に関する内閣府令（以下、「取引業内閣府令」とする）125 条 1 項 12 号によると、「取引所金融商品市場における上場金融商品等又は店頭売買有価証券市場における店頭売買有価証券の相場若しくは相場若しくは取引高に基づいて算出した数値を変動させ、若しくはくぎ付けし、固定し、若しくは安定させ、又は取引高を増加させることにより実勢を反映しない作為的なものを形成させるべき当該上場金融商品等若しくは当該店頭売買有価証券に係る買付け若しくは売付け若しくはデリバティブ取引又はこれらの申込み若しくは委託等若しくは受託等をする行為を防止するための売買管理が十分でないと認められる状況」である。また、高速取引行為者についても同種

[317] 斉藤ほか・前掲（注 306）201 頁（平山）。

の規制が課せられる（金商法 66 条の 57)。すなわち、高速取引行為者は業務の運営の状況が、「業務の運営の状況が公益に反し、又は投資者の保護に支障を生ずるおそれがあるものとして内閣府令で定める状況にあること」にならないようにその業務を行わなければならない。ここでいう「状況」については、金融商品取引業等に関する内閣府令 235 条 2 号によると、「取引所金融商品市場における上場金融商品等の相場若しくは相場若しくは取引高に基づいて算出した数値を変動させ、若しくはくぎ付けし、固定し、若しくは安定させ、又は取引高を増加させることにより実勢を反映しない作為的なものを形成させるべき当該上場金融商品等に係る買付け若しくは売付け若しくはデリバティブ取引又はこれらの申込み若しくは委託等をする行為を防止するための売買管理が十分でないと認められる状況」である。ただし、これら規定に対応する刑事罰規定や課徴金規定は存在しないため、これら業者規制をつうじた、AI・アルゴリズム投資に対し相場操縦エンフォースメントをして事前に防ぐということはできない。

以上のこれらの規制状況を考慮すると、AI・アルゴリズムを利用した投資において、自然の需給関係に反する変動取引がなされた場合は、なお金商法 159 条 2 項における相場操縦規制の対象とならない可能性が高いことになる。そこで、不公正取引規制の一般条項である金商法 157 条の適用可能性をその法的性質・要件解釈にさかのぼって検討する。

第 4 目　一般条項の適用可否

（1）法的性質

証券市場に対する投資家の信頼を明らかに害する行為であっても、相場操縦型、もしくはインサイダー取引型のいずれかの禁止規定で捕捉するのが困難な場合が存在する。そのような行為については、不正行為を包括的に禁止する本規定で捕捉する。金商法 157 条 1 項は、何人も「有価証券の売買その他の取引又はデリバティブ取引等について、不正の手段、計画又は技巧をすること」をすることを禁止している。本規制に違反した場合は、課徴金は課されず、刑事罰のみが科される（金商法 197 条 1 項 5 号)。そのため、犯罪構成要件として「不正の手段、計画又は技巧」の意義を明確にする必要がある。

（2）要件解釈

この要件について、判例は以下のような枠組を示している。最決昭和 40 年 5 月 25 日（刑集 155 号 831 頁）那須硫黄工業事件によると、「『不正の手段』とは、有価証券の取引に限定して、それに関し、社会通念上不正と認められる一切の手段をいうのであつて、文理上その意味は明確であり、それ自体において、犯罪の構成要件を明らかにしていると認められる（第一審判決の確定した事実によれば、本件は、被告人が、無価値の株券に偽装の株価をつけるため、証券会社の外務員二名と共謀の上、同人らをして、判示会社の株式につき、権利の移転を目的としない仮装の売買を行わせたというのであり、かような行為が、証券取引法五八条一号（現行金商法 157 条 1 項―筆者注）にいわゆる『不正の手段』に該当することは

明白である)」[318]とした。この事例は、被告人に誤認目的を認定することはできず、仮装売買型の相場操縦規制（当時の証券取引法 125 条 1 号型）を適用することができなかったものであった。

なお、この事件の控訴審（東京高裁昭和 38 年 7 月 10 日下刑集 5 巻 7・8 号 651 頁）では、「証券取引法第五八条第一号にいう『不正の手段』とは、取引所取引たると、店頭取引たるとを問わず、有価証券の売買その他の取引について、詐欺的行為、すなわち、人を錯誤におとし入れることによつて、自ら、または他人の利益を計ろうとすることであると解するを相当とする（下線筆者)」と、不正性として詐欺的行為の要件を付け加えている[319]。ただし、本条が適用されたケースは極めて稀であり、公刊物に搭載されているものは現在、上記の那須硫黄工業株事件の 1 件のみである[320]。

この要件について、①この規制の母法である米国の SEC 規則 10b-5 の射程範囲が判例の積み重ね通じて拡大されてきたのは、1933 年以前から発展していた民事上の「詐欺(fraud)」の概念が基礎とされたもので[321]、その条文にも”To employ any device, scheme, or artifice **to defraud**”とあるため、詐欺的行為に限定して解釈すべきであり、157 条 2 項・3 項には具体的行為として詐欺的行為が列挙されていることから、157 条 1 項も同様に詐欺的行為に限定する見解、②条文の文言上、「詐欺的」等の要件がない以上これに限定せず、むしろ、本罪をインサイダー取引や相場操縦の要件を厳密には充足しないものの、それらと同様の当罰性を有する行為に積極的に適用すべきであるという見解[322]が存在する。

しかし、この条文の構成要件の抽象性により、明確性の原則に則れば、そもそもこの条文で刑事罰を科すにはかなり困難であるものと考えられる上に、その上法定刑も金商法では最も重い罰則であるため、なおさら②の見解を支持することはできない。加えて、元来この条文の母法である SEC 規制 10b-5 では、この条文でインサイダー取引を摘発していたところ、明確性の原則に反する可能性が高いという理由で日本では同様の運用ができず、そのため、1988 年証券取引法改正で新たにインサイダー取引規制が設けられたという背景がある[323]。かりに 157 条 1 号を適用するにしても、①の見解を支持すべきであろう。そうすると、AI・アルゴリズムの利用者に対し詐欺的行為を認定できるか否かについての問題となる。しかし、AI・アルゴリズムの利用者には上記の検討から主観的要素（この場合は詐欺の目的）を認定することは困難である。そのため、一般条項による解決を図るのではなく、もっぱら AI・アルゴリズムによる相場操縦行為に対する利用者の規制は立法的解決が望ましいように思われる。

---

318 この見解は東京地判平成 10 年 5 月 14 日判時 1650 号 145 頁も同旨。

319 東京高裁平成 7 年 9 月 26 日判時 1549 号 11 頁も同旨。

320 山口厚『経済刑法』（商事法務、2012 年）211 頁。

321 松尾直彦『金融商品取引法 [第 2 版]』（商事法務、2020 年）592 頁。

322 神田ほか編・前掲（注 305）8 頁（藤田)。

323 神田ほか編・前掲（注 305）110 頁（神作)。

第４項　立法的解決

以上より、開発者や販売者に対しては、自然の需給に基づかない相場を作出するような取引が行われないように AI・アルゴリズムを構築・管理することの義務付け、それに伴う罰則規定の創設が望ましいと考える。ここで注意すべきなのは、かりに AI・アルゴリズムが相場操縦行為に該当する取引を実行したと認められる場合にはじめて、証券取引等監視委員会などの専門機関による改善措置命令を出すことを可能にし、それでもなお従わない場合に刑罰規定・課徴金規定を付すことで、AI・アルゴリズム投資における相場操縦規制の実効性をバランス良く担保することができるということである。

また、利用者（高速取引行為者）については、現行の規制に対する課徴金・罰則規定の追加が望ましいものと思われる。とりわけ、高速取引行為者の業者規制にこれらを追加すべきであると思われるが、その際、製造者に対する規制と同様に、AI・アルゴリズムによる相場操縦行為が確認されれば直ちに摘発するのではなく、専門機関の改善措置を間にはさみ、それでもなお、これに従わない場合にはじめて摘発するという形式がよいだろう。

このことは、AI・アルゴリズム投資の濫用、すなわち、主観的要件認定の困難さを理由に本来であれば処罰されるべき相場操縦行為者が責任から免れようとする態度による処罰の間隙を埋めるのみならず、適正な方法で AI・アルゴリズムを利用する利用者や、開発設計を行う製造者の処罰範囲を明確にする効果も期待できる。

第５項　小括

以上の検討から、自動的に取引を行うアルゴリズム・AI について、自然の需給に基づかない相場を作出するような取引が行われないようにアルゴリズム・AI を構築・管理することを開発者や販売者に義務付け、また、そうした管理義務に違反する場合には、課徴金を課したり、刑事罰を科したりする規定の導入を早急に考えるべきである。

その理由は以下の通りである。現行規定の解釈枠組では、仮装売買、見せ玉、変動取引のような客観的事実から相場操縦規制における主観的要件が認定される。しかし AI・アルゴリズムが相場操縦をする可能性を利用者が認識することをもって相場操縦規制の主観的要件を充足させるべきでない。それは処罰範囲の拡大を意味するだけでなく、AI・アルゴリズム投資の便益を損なうことにつながるからである。立法的解決の根拠は、主観的要件の認定の困難さによる処罰の間隙への対応のみならず、客観的事実からの主観的要件の推認の範囲を制限することにあり、専門機関の改善措置を間にはさみ、それでもなお、これに従わない場合にはじめて摘発するという形式が望ましい。

第２款　インサイダー取引

第 1 項　問題の所在

近時、AI・アルゴリズムを用いて投資運用を補助するロボアドバイザー[324]や、「AI 株式ポートフォリオ診断」[325]の開発が進められているものの、その利便性の反面、インサイダー取引の禁止（金商法 166 条以下）規制に該当しうることが指摘されている。本項では、上記のうち、AI・アルゴリズムが介在する証券取引において、インサイダー取引規制がどのように関わるかを検討した先行研究[326]に関しその妥当性を検証する。

第 2 項　インサイダー取引規制の概説

金商法 166・167 条は、上場会社や公開買付け等を行う者と特別の関係にある者が、未公表の重要事実を知って、その事実が公表される前に一定の有価証券等を売買等することを禁止する。

インサイダー取引の規制根拠は、投資者間の不公平そのものを根拠とする訳ではなく、そのような取引が行われることが、証券市場に対する信頼を失わせ、その健全な発展の障害になることにあるとされる[327]。換言すれば以下の通りである。すなわち、特別の立場にあることにより情報を知り得る者が、情報を知らない一般投資家の犠牲のもとに利益を得ることが許されるような証券市場は、必ずしも公正なものとはいえない。証券市場の公正さが失われ、市場に対する信頼を失った投資家は、市場から遠ざかってゆく。そうすると、企業・投資家は、証券市場で資金調達をすることが困難となり、証券市場を通じた資金分配という制度も維持できなくなり、証券市場の健全な発展が害されてしまうと考えられている[328]。まとめると、①投資家間の情報の非対称性に起因する不公平を是正して公平を確保することにより投資者保護を図ること、②証券市場における価格形成の公正性を確保すること、そして③ 投資者の市場参加による流動性確保を通じた証券市場の効率性を確保することである[329]。

本条は、1988 年改正証券取引法において導入された。もっともインサイダー取引そのも

---

[324] ロボアドバイザーとは、AI を利用して投資家の代わりに資産運用のアドバイスや運用の補助をするサービスである。具体的には、「アドバイス型」と「投資一任型」のものがあり、前者は利用者のリスク許容度に応じて、最適な資産の組合せ（ポートフォリオ）の提案を行う。また、後者では、利用者が運用資金を入金するだけで、ポートフォリオに応じて自動で買い付けが行われる。

[325] SMBC 日興証券株式会社が 2020 年より提供するサービスである。追加投資金額、購入を検討している銘柄とその市場、リスク許容度を選ぶと、AI がリスクと期待収益を考慮して株式投資のポートフォリオを提案し、保有している銘柄の情報を基に、AI が売買すべき銘柄を表示して資産運用をサポートする。

[326] アルゴリズム・AI の利用を巡る法律問題研究会・前掲（注 4）1 頁以下。なお、本節でもこの論文の内容を単に「先行研究」と呼ぶことがある。

[327] 斉藤ほか・前掲（注 306）203 頁（平山）。

[328] 斉藤ほか・前掲（注 306）203 頁（平山）。

[329] 松尾・前掲（注 321）623 頁。

のは、不公正取引の一般的禁止条項たる旧証券取引法 58 条 1 項（現・金商法 157 項 1 項）によって把握されるとされていたが、前述のように、金商法 157 条 1 項を理由に刑事罰を科すには、規定の仕方が抽象的なために明確性の原理の観点から適用が困難であったこと、そして、そもそもインサイダー取引がそれほど悪質な犯罪とは一般に認識されていなかったこと[330]から、同規定によってインサイダー取引が処罰された事件は存在しなかった[331]。

しかし、タテホ化学工業が債券先物取引によって約 280 億円を超える損失を発生させたところ、これが公表される前に、タテホ化学工業の取締役や取引銀行の阪神相互銀行が所有していたタテホ工業の株を売却して損失を免れた、という 1987 年 8 月のタテホ化学工業事件を機に、立件こそされなかったものの、インサイダー取引に対する社会的な非難が高まるに至った[332]。この事件を直接の契機として、インサイダー取引に関する規制が新設された。

インサイダー取引規制は、大きく会社関係者等によるものと、公開買付者等関係者等によるものに分かれる。例えば、会社関係者等によるインサイダー取引規制の対象となる主体は、会社関係者と情報受領者に分けられる。会社関係者として規制対象となる主体は、金商法 166 条 1 項各号に規定されており、例えば、上場会社等の役員、代理人、使用人その他の従業者（以下、「役員等」とする）は、当該上場会社等にかかわる業務等に関する重要事実を、その者の職務に関し知ったときは、当該業務等に関する重要事実の公表がされた後でなければ、当該上場会社等の特定有価証券等にかかわる売買等をしてはならない、とする。また、金商法 166 条 3 項は、会社関係者から重要事実の伝達を受けた者または職務上当該伝達を受けた者が所属する法人の他の役員等であって、その者の職務に関し当該重要事実を知った者（以下、「第一次情報受領者」という。）も、当該重要事実の公表がされた後でなければ、当該上場会社等の特定有価証券等にかかわる売買等をしてはならないとしている。

本罪の法効果は、現行法においては 5 年以下の懲役もしくは 500 万円以下の罰金（またはこれらの併科）となっている（金商法 197 条の 2 第 13 号）。また、両罰規定によって、法人には 5 億円以下の罰金が科される（金商法 207 条 1 項 2 号）。不法収益の剥奪に関しても、インサイダー　取引によって得た財産は必要的没収・追徴の対象となっている（金商法 198 条の 2）。さらに、自己の計算によってインサイダー取引を行った者、さらに顧客の計算によりインサイダー取引を行った金融商品取引業者についても、課徴金が課されることになる（金商法 175 条）。

第 3 項 AI・アルゴリズムとインサイダー取引

本報告では、AI・アルゴリズムを利用して法人が取引を行う際に、いかなる場合に法人ない

---

330 佐伯仁志「インサイダー取引」西田典之編『金融業務と刑事法』（有斐閣、1997 年）220 頁参照。

331 斉藤ほか・前掲（注 306）204 頁（平山）、山口・前掲（注 320）229 頁、神田ほか・前掲（注 308）109 頁（神作）。

332 斉藤ほか・前掲（注 306）204 頁（平山）。

しはその利用者が処罰され、もしくは課徴金が課されるのかを以下の事例[333]で検討する。

第 1 目　想定事例

ある上場企業 A の役員 X が、その職務に関し、A 社業務に関する未公表重要事実（金商法 166 条 2 項に該当するいずれかの事実）を知り、その後、企業 B の役員 Y は、X から当該未公表重要事実の伝達を受けた。B 社は、AI・アルゴリズムを利用して、企業内のポートフォリオの構築・改善を行っていた。しかし、未公表重要事実の伝達を受けた Y は、当該ポートフォリオの構築・改善業務とは無関係であり、この業務の責任者は A 社従業員の Z であった。Y は B 社内のデータベースに当該未公表重要事実を入力した。
①Z は当該未公表重要事実の伝達を Y から受け、Y が B 社のポートフォリオにこれを入力したことを知っていた。この状況のもと、当該入力に基づいて B 社の AI・アルゴリズムの判断で A 社株式の購入が行われた。 ②Z はこの未公表重要事実が入力されたことを知らなかった。この状況の下、当該入力に基づいて B 社の AI・アルゴリズムの判断で A 社株式の購入が行われた。

第 2 目　【①事例】の検討

（1）金商法 163 条 1 項・3 項の要件解釈

①事例において Y・Z はインサイダー取引規定のいかなる主体に該当するか。AI・アルゴリズムは法人ではないため、この事例において金商法 163 条 3 項前段や後段に AI・アルゴリズムは該当しない。そうすると、Z は「職務上当該伝達を受けた者（B 社役員 Y）が所属する法人（B 社）の他の役員等であって、その者の職務に関し当該業務等に関する重要事実を知ったもの」に該当し、当該上場会社等（A 社）の株式を重要事実公表前に購入した主体のように見える。この場合、Y が情報提供者、Z が第一次情報受領者となる。

次に「重要事実を『知って』」の解釈については、一般投資家であれば投資判断に著しい影響を及ぼすに足りると認識する事実であるという確定的な認識のみならず未必的な認識にすぎない場合も含まれるとする[334]。ただし、重要事実を「知った」ことと「売買等」の間における因果関係の存在は要件とされていないので[335]、未公表重要事実を知る前から A 社株式の取引を行っていたとしても、これを知ったあとに AI・アルゴリズムにより自動的になされた取引は金商法 166 条 3 項の要件に該当する。しかしこの規制では、未公表重要事実を知っていて、それを利用していない場合だとしても常にインサイダー取引違反（金商法 166 条 3 項）であるとするならば不必要に規制範囲が拡大するおそれがある。そこで、内部

333 アルゴリズム・AI の利用を巡る法律問題研究会・前掲（注 4）28 頁の事例を参考にしている。

334 東京高判平成 29 年 6 月 7 日 裁判所ウェブサイト（原審:東京地判平成 29 年 1 月 13 日判例タイムズ 1449 号 166 頁。他に、木目田裕監修・直村あさひ法律事務所・危機管理グループ編『インサイダー取引規制の実務（第 2 版）』（商事法務、2014 年）236 頁（小林）。

335 松尾・前掲（注 321）645 頁。

者取引規制の対象に含まれるものの、証券市場の公正性・信頼性を害しないような取引については、類型的に適用を除外する規定が金商法 166 条 6 項の適用除外要件に定められている。

インサイダー取引規制の構成要件は、取引の実質的な不正を要件とすることなく、形式犯として定められていることから、インサイダー取引規制の趣旨である証券市場の公正性・健全性に対する投資者の信頼の確保の観点から、類型的に規制対象とする必要のないと考えられる取引が具体的に列挙され[336]、そのうち上記事例に関連するのは金商法 166 条 6 項 12 号類型である。そこでは、①上場会社等に係る第 1 項に規定する業務等に関する重要事実を知る前に締結された当該もしくは上場会社等の特定有価証券等に係る売買等に関する契約の履行、②上場会社等に係る同項に規定する業務等に関する重要事実を知る前に決定された当該上場会社等の特定有価証券等に係る売買等の計画の実行としての売買等、③その他これに準ずる特別の事情に基づく売買等であることが明らかな売買等が規定されている。

ただし、どの類型が内閣府令によって限定を受けるのかについては①・②・③の全ての類型で限定を受けるのか、①・②のみの類型で限定を受けるのか、③のみ類型で限定を受けるのかで争いがある[337]。この点、有価証券の取引等の規制に関する内閣府令 59 条 1 項では、上記①・②の類型に関する規定が存在する一方で、③に関する規定は存在していない。この点、③類型に対応する内閣府令は存在しないため③による適用除外は存在しないとする見解もあるが[338]、重要事実を知ったことと無関係に行われる取引であることが客観的に明らかな場合については適用除外されると解すべきであるとされる[339]。これは、最大判平成 14 年 2 月 13 日民集 56 巻 2 号 33 頁[340]の趣旨や立案担当者の趣旨[341]にも対応する。

---

336 例えば、新株予約権の行使による株券の取得（2 号）、防戦買い（4 号）、安定操作取引（5 号）、重要情報を知る者同士の取引（7 号）、重要事実を知る前に締結された契約の履行（12 号）である。詳しくは松尾・前掲（注 321）656 頁を参照。

337 岩原紳作ほか「金融商品取引法セミナー（第 17 回）追補：内部者取引規制と公開買付規制」ジュリスト 1417 号（2011 年）105 頁～107 頁。

338 服部秀一『インサイダー取引規制のすべて』(商事法務研究会、2001 年) 222 頁。

339 松尾・前掲（注 321）665 頁。

340 本判決は、上場会社等の役員等の短期売買利益の返還を定めた 164 条に関するものである。判示によれば「同条（証券取引法 164 条―筆者注）8 項は、取引の態様等を勘案してこのような秘密の不当利用の余地がないものと観念される取引の類型を定めることを内閣府令に委任したものであるが，上記の目的を達成するために同条 1 項の規定を適用する必要のない取引は内閣府令で定められた場合に尽きるものではなく，類型的にみて取引の態様自体から 上記秘密を不当に利用することが認められない場合には，同項の規定は適用されないと解するのが相当である」という。ここで留意すべきなのは、単に重要情報を「知った」のみならず、それを「利用する」ことに言及されていることである。

341 横畠裕介『逐条解説インサイダー取引規制と罰則』(商事法務研究、1989 年) 9～10 頁参照。そこでは、重要事実を知った会社関係者等が上場株券等の売買に当たってその情報を「利用した」ものでないとか、当該情報に「基づいて」取引をしたのではないなど、内心の意思や動機を問題としているのではなく、そのような重要事実を知ったことと無関係に行われる売買等であることが明確であるような特別な事

有価証券の取引等の規制に関する内閣府令（以下、「取引府令」とする）59 条 1 項で、上記①に対応する規定としては、1 号類型が挙げられる。同規定によると、重要事実を知る前に上場会社等との間で書面（第 13 条 5 項に規定する電磁的記録を含む）による契約をした者が当該契約の履行として、当該書面に定められた当該売買等を行うべき期日又は当該書面に定められた当該売買等を行うべき期限の十日前から当該期限までの間において売買等をする（1 号）から始まり、包括的な適用除外規定（14 号）に至る。

この 14 号の包括規定では、①業務等に関する重要事実を知る前に締結された特定有価証券等に係る売買等に関する書面による契約の履行又は業務等に関する重要事実を知る前に決定された特定有価証券等に係る売買等の書面による計画の実行として売買等を行うこと、②業務等に関する重要事実を知る前に、(1) 当該契約若しくは計画又はこれらの写しが、金融商品取引業者に対して提出され、当該提出の日付について当該金融商品取引業者による確認を受ける、(2) 当該契約又は計画に確定日付が付される、(3) 当該契約又は計画が第 166 条第 4 項に定める公表の措置に準じ公衆の縦覧に供される、のいずれかがなされたこと、③当該契約の履行又は当該計画の実行として行う売買等につき、売買等の別、銘柄及び期日並びに当該期日における売買等の総額又は数が、当該契約若しくは計画において特定されていること、又は当該契約若しくは計画においてあらかじめ定められた裁量の余地がない方式により決定されること、のすべてを満たす場合に適用除外となると定められる。2020 年金商法改正により、このような「知る前」契約においては電磁的記録による作成のものも対象内となったが、1 号類型では電磁的記録によるものを含む A 社と Z 間での書面による契約がないため適用できないので、14 号類型では、Z が重要事実を「知る前」に AI・アルゴリズムによって構築されたポートフォリオに基づいて取引が実行されたことが証明され、14 号所定の要件に従い、契約にかかる確認を受ける、確定日付が付されるもしくは公衆の縦覧に供される限りで適用除外となる。

（２）利用者の積極的作為義務の有無に関する検討

①事例の Z は未公開重要事実を知った後に、A 社株式の売買を停止する措置を AI・アルゴリズムに対して行う義務を有するかという問題が、信託契約や投資一任契約を利用した場合との比較で挙げられている[342]。

まず、信託契約や投資一任契約は金商法 166 条 1 項における「売買等」に該当するのかという問題から出発する。166 条 1 項における「売買等」とは、株券等について有償でその所有権を移転することをいうと一般に解されており、売買のほか、交換、代物弁済や現物出資なども「売買等」に当たる一方，相続や贈与による取得は該当しないと解される[343]。他人に

情があるという客観的な状況が存在することを要件とするものであり、極めて限定された場合にのみ該当するものとする。その例として、担保株式の処分として売却決定者の裁量の余地なく機械的に売却する場合のように、他人の指示を受けて機械的に売買等に関与するケースが挙げられる。

342 アルゴリズム・AI の利用を巡る法律問題研究会・前掲（注 4）32 頁。

343 神田ほか・前掲（注 308）122 頁（神作）。

売買等の委託、指図することも 166 条 1 項における「売買等」に含まれるため、信託方式等であっても、役員等が重要事実を知って信託契約等を締結・変更するのであれば、その役員等についてインサイダー取引（166 条 3 項）が成立することになる[344]。

次に重要事実を知らずに信託契約等を締結した場合を考察する。上場会社が信託方式又は投資一任方式によって自己株式取得を行う場合、実際には第三者である信託銀行等が買付主体となるため、会社関係者が重要事実を知って売買等を行う場合に該当するかどうかが問題となる。この問題に関し、「インサイダー取引規制に関するＱ＆Ａ」（金融庁）によると[345]、信託契約又は投資一任契約の締結・変更が、当該上場会社により重要事実を知ることなく行われたものであって、当該上場会社が契約締結後に注文に係る指示を行わない形の契約である場合、もしくは、当該上場会社が契約締結後に注文に係る指示を行う場合であっても、指示を行う部署が重要事実から遮断され、かつ、当該部署が重要事実を知っている者から独立して指示を行っているなど、その時点において、重要事実に基づいて指示が行われていないと認められる場合は、一般に上記の会社関係者が重要事実を知って売買等を行う場合に該当しないと考えられる。

この関連で、信託契約等で買付けを中止できる契約になっている場合、積極的に中止の措置をとらなければ規制違反になるのではないかという問題も提起される。しかし、重要事実を知る前に締結された信託契約等に基づいて信託銀行等が売買等をするのであれば、重要事実を知った役員等の裁量の余地はなく、たとえ信託契約等に基づき、それ以降も継続的に売買等がなされたとしても、特に一般投資者に比べて有利な投資判断が行われることにはならず、投資者の市場に対する信頼を損なうことにはならないとされる[346]。

（3）当てはめ

AI・アルゴリズムに何らの修正も行わず、未公表重要事実が与えられることもない場合は、信託契約や投資一任契約を利用して取引を行う場合であって、契約締結後に指示を行わない場合や重要事実に基づいて指示が行われていないと認められる場合と同様に、インサイダー取引規制違反とならないと考えられる[347]。

---

344 金融商品取引法研究会「インサイダー取引規制と自己株式」（2015 年）8 頁。

345 金融庁 証券取引等監視委員会「インサイダー取引規制に関するＱ＆Ａ」（2019 年）12 頁。なお、166 条 6 項 12 号の適用除外規定に求める見解も存在するが、そもそもこの規定が不明確なものであり、広く重要事実を知ったことと無関係に行われる売買等であることが明らかなものを適用除外にする規定であるというのが金融庁の伝統的解釈であるから、166 条 1 項所定の重要事実を知った者の売買に該当しないとして、これにより規制違反にならないと理解すべきとする。

346 金融商品取引法研究会『インサイダー取引規制と自己株式』（2015 年）8~9 頁。この事例においては、当該中止義務を不真正不作為犯における作為義務と同一視し、その場合の先行行為を信託契約（投資一任契約）に求めるものとする。しかし、当該信託契約では以後のインサイダー取引が発生する何か切迫した危険があるということがあるとは言えないため、この先行行為に基づく作為義務、作為との同質性というのは認められないという（37 頁以下参照）。

347 アルゴリズム・AI の利用を巡る法律問題委員会・前掲（注 4）33 頁。

事例①の取引責任者 Z は、未公表重要事実を知っており、現に AI・アルゴリズムによる取引を継続したり中止したりすることができる状況下であえて取引を中止せずに継続しているので、信託契約と比較して重要事実を知った役員等の裁量の余地があるから、適用除外要件（166 条 6 項 12 号）に該当しない限り、重要事実に基づいて指示を行う場合と同様に 166 条 3 項の罪が成立しうる。この場合、Y は Z の共犯となりうる。

しかし、外形的に Y や Z のような利用者の行為がインサイダー取引に該当し、刑罰や課徴金といった制裁の負担を負うとするならば、AI・アルゴリズムによるポートフォリオ構築の普及を阻害しかねないし、利便性を損なうことにもなりうる。そこで、当該事例において「知る前」契約の適用除外（金商法 166 条 6 項 12 号、取引府令 59 条 1 項 14 号）を適用して利用者らの負担を軽減するような体制が望ましい。これを実現するには、取引府令 59 条 1 項 14 号の要件を充足しうるようなシステム構築が必要である。具体的には、利用者が金融商品取引業者ならば、取引府令 59 条 1 項 14 号ロ(2)の要件に従い確定日付が付されることが、そうでない利用者ならば、取引府令 59 条 1 項 14 号ロ(1)の要件に従い対象となる有価証券等の取引相手型となる金融商品取引業者の確認を受けること、もしくは 14 号ロ(3)の要件に従い、公衆の縦覧に供されることが可能となるようなシステム構築が製造者には求められる。

第 3 目　【②事例】の検討

（1）先行研究の検討スキーム

先行研究では、Z の行為の金商法 166 条 3 項該当性、B 社に対する課徴金（金商法 175 条)、氏名公表措置（金商法 192 条の 2)、業者規制（金商法 40 条）の適用を検討する。以下、それぞれについてみていくことにする。

まず、Z の行為の 166 条 3 項該当性について、②事例では自然人や法人ではない「AI・アルゴリズムが未公表重要事実を知って取引し」、Z は重要事実を「知って」取引しているわけではないように見えるので、Z は 166 条 3 項の構成要件に該当しない。このとき、責任を負うべき人間の主体が存在しないということになるため、B 社に対する責任を検討することになる。

そこで考えられるのが、B 社に対する課徴金の可能性（175 条）である。法人の役員等が当該法人の計算でインサイダー取引を行った場合は、当該法人が課徴金納付命令の対象となるので、②事例では重要情報の第一次情報受領者を B 社とすることも可能である。この課徴金納付命令については、行政上の措置であることから、一般に、法人に対して課徴金納付命令を行う際には行為者を特定する必要はないと解されている。しかし、インサイダー取引規制（166 条 1 項・3 項）では「知った」ことが主観的構成要件要素とされているため、法人に対して課徴金納付命令をするためには、行為者を特定してその者が重要事実を「知った」ことが認定される必要があるが、Z は重要事実を知らないため解釈上この条文は適用できないという。

次に考えられるのが、氏名公表措置（金商法 192 条の 2）や業者規制（金商法 40 条 2 号）である。前者について金商法 192 条の 2 では、「内閣総理大臣は、公益又は投資者保護のため必要かつ適当であると認めるときは、内閣府令で定めるところにより、この法律又はこの法律に基づく命令に違反する行為…を行つた者の氏名その他法令違反行為による被害の発生若しくは拡大を防止し、又は取引の公正を確保するために必要な事項を一般に公表することができる」と規定する。後者について、金融商品取引業者ならば、取引業内閣府令 123 条 1 項 5 号に規定される「その取り扱う法人関係情報に関する管理又は顧客の有価証券の売買その他の取引等に関する管理について法人関係情報[348]に係る不公正な取引の防止を図るために必要かつ適切な措置を講じていないと認められる状況」にならないようにする義務が金商法 40 条 2 号をつうじて課され、高速取引業者ならば、取引業府令 337 条 1 号に規定される 123 条 1 項 5 号所定の状況と同様にならないようにする義務が金商法 66 条の 57 第 2 号をつうじて課される。これらのエンフォースメント手段として、業務改善命令（金融商品取引業者：金商法 51 条、高速取引業者：金商法 66 条の 67）や許可取消処分・業務停止命令（金融商品取引業者：金商法 52 条、高速取引業者：金商法 66 条の 68）と氏名公表措置[349]（金商法 192 条の 2）が考慮されるにとどまる。そうすると、未公表重要事実が与えられた AI・アルゴリズムを利用した取引の場合、それを利用しない取引の場合との法効果に差異が生じることになりうるし、自己のポートフォリオ構築のために AI・アルゴリズムを利用する法人が金商法上の規制対象とならない場合はこの業法規定を適用することはできない。せいぜい、対象者の社会的信用等に不利な影響を与えるものであり，法令違反行為に対する制裁としての性質を有する氏名公表措置が適用できるにすぎないのである[350]。

（2）現行法におけるありうる対応とその評価

例えば金商法 167 条の 2（情報伝達・取引推奨規制）において、第 1 次情報受領者にとって未公表重要事実の利用のみをもって、直ちに規制対象とはされていないことから、法人の役員等が未公表重要事実を知らない場合には、アルゴリズム・AI を通じて間接的に未公表重要事実を利用した場合であっても、規制対象とする必要はないのではないか。しかし、こうした場合が規制の対象とならないとすると、AI・アルゴリズムに未公表重要事実が与えられる仕組みを故意に構築する余地を生みかねず、著しく不公平な取引を促進することになりかねない。

---

348 この法人関係情報とは、金融商品取引業等に関する内閣府令 1 条 4 項 14 号に規定されている。その内容として、金商法 163 条第 1 項に規定する上場会社等の運営、業務又は財産に関する公表されていない重要な情報であって顧客の投資判断に影響を及ぼすと認められるもの並びに金商法 27 条の 2 第 1 項に規定する公開買付け、これに準ずる株券等の買集め及び金商法 27 条の 22 の 2 第 1 項に規定する公開買付けの実施又は中止の決定（金商法第 167 条第 2 項ただし書に規定する基準に該当するものを除く。）に係る公表されていない情報をいう。

349 氏名公表措置の適用可能性は先行研究では触れられていない。

350 木目田・上島・前掲（注 334）607 頁（尾崎・有松）。

そこで、現行法の対応として先行研究では以下の可能性を示唆している。

まず、行為者を法人（B 社）として、「B 社が未公表重要事実の伝達を受けて（＝知って）取引を受けた」と解し、課徴金は法人に対しても課すことができるところ、この見解で課徴金を課すことができるのではないかという。しかしインサイダー取引規制で、法人に対して課徴金納付命令（175 条）をするためには、行為者を特定してその者が重要事実を「知った」ことが認定される必要があるから、行為者を特定しなければならず、先行研究のスキーム上では誰も該当しないことになる。次に、一般規定である金商法 157 条 1 号の適用可能性を模索するが、第 2 項第 3 目でも指摘したように、金商法 157 条 1 項を理由に刑事罰を科すには、規定の仕方が抽象的なために明確性の原理の観点から適用が困難であったことなど、インサイダー取引規制が導入された背景をも考慮すると、この条文を適用して AI・アルゴリズムに未公表重要事実が与えられる仕組みを故意に構築した者に刑事罰を科すことは困難である。

そして、上記二つの提言とはやや方向性が異なるが、先行研究では過失によって未公開重要事実が AI・アルゴリズムに与えられた場合[351]でも規制の必要性を示唆している。そこでは、当該法人に著しく不公平な状況が生じることを防ぐ必要があると考える場合、当該 AI・アルゴリズムに未公表重要事実が与えられたことをもって、法人がこれの伝達を受けたものとして、課徴金賦課の対象となるインサイダー取引規制違反に問うことも考えられるという。確かに、インサイダー取引規制の理解不足から未公表の重要事実等があるにもかかわらず自己株取得を行ってしまったいわゆる「うっかりインサイダー」と呼ばれる不注意事案や、課徴金額がわずか数万円の軽微事案であっても、網羅的に課徴金調査、勧告、課徴金納付命令の発出がなされてきた[352]。しかし、このような硬直的な運用では，行為の態様・悪質性と制裁の内容とがバランスを失し、課徴金制裁を課すことが必ずしも妥当でない場合があり、「不当利得の没収という前提を維持しつつも，制裁色を強めている…行政手続だから構成要件に形式的に該当していれば納めてもらいます、という姿勢がもたらす市場への副作用は測り知れない。法律の建付けはどうあれ、やった行為に応じて過不足なく責任を取ってもらう手段として課徴金を認識することにより，私たちの仕事の水準も向上する」[353]、「現在の体制においては，いわゆる“うっかりインサイダー”というようなものは摘発しないという運用が確立しております」[354]と実務的見解も過失によるインサイダー取引は摘発すべきでないということが窺える。以上の検討から、過失による AI・アルゴリズムによるインサイダー取引規制の適用を拡張する必要はないように思われる。

---

[351] AI・アルゴリズムに未公表重要事実が与えられない仕組みとすることが意図されていたが、システムの不備もしくは役員等の理解不足により AI・アルゴリズムに未公表重要事実が意図せず与えられてしまったような場合が想定されている（アルゴリズム・AI の利用を巡る法律問題研究会・前掲（注 4）35 頁参照）。

[352] 木目田・上島・前掲（注 334）614 頁。

[353] 大森泰人「課徴金(下)」金法 1896 号（2010 年）6 頁以下。

[354] 大証金融商品取引法研究会報告「市場監視の実際」（2010 年）15 頁。

（3）Y の行為の再検討

先行研究では専ら Z が取引主体であることを強調して、Y の行為の検討を行ってはいない。しかし、Y の行為は金商法 166 条 3 項の構成要件に該当するかについての検討は必要である。

Y は金商法 166 条 3 項前段の、「会社関係者（A 社役員 X）から当該会社関係者が第一項各号に定めるところにより知つた同項に規定する業務等に関する重要事実の伝達を受けた者」に該当する。ただし、Y は A 社の役員等ではないため、金商法 167 条の 2（情報伝達・取引推奨規制）の対象者には該当しない。

そこで問題となるのが、Y が「当該上場会社（A 社）の特定有価証券等に係る売買等を」したといえるかである。事例②では、未公開重要事実を知らない Z（ないしは AI・アルゴリズム）が「当該上場会社（A 社）の特定有価証券等に係る売買等を」している。このとき当該取引の利益帰属主体は B 社であるから、その B 社を代表する Y・Z が行為者と考えられるところ、「Y は未公開重要事実の存在を利用して、情を知らない Z（ないしは AI・アルゴリズム）に A 社株式を売買させる」という構成が可能である。この場合、Y のみが刑罰・課徴金として制裁を受けることになるだろう。

（4）適用除外規定に関する検討

金商法 166 条 1 項・3 項においては、「重要情報を知った」ことと「有価証券等の取引を行った」こととの因果関係が要求されていないため、規制対象が広範になることから、適用除外規定（166 条 6 項）により規制対象とならない取引形態を規定したことは先にも述べたとおりである。その場合、以下の事例に適用除外規定の射程が及ぶかどうかが問題となる。

【③事例】Y は当該未公開重要事実を知った上で B 社の AI・アルゴリズムに当該未公開重要事実を入力し、A 社株式が取引された（取引責任者は Z とする）。しかし、取引実行時の AI・アルゴリズムと、その未公開重要事実のみが与えられていない AI・アルゴリズムを再現したところ、全く同一の判断をしたことが判明した。

【④事例】取引責任者を兼任する Y は当該未公開重要事実を知っていたが、これを B 社の AI・アルゴリズムに与えることはなかった。その後、当該 AI・アルゴリズムの判断に基づき A 社株式の取引が実行された。この AI・アルゴリズムは学習を行うものであったが、この投資判断基準プロセスを証明することはできなかった。

③事例では、未公開重要事実の入力と取引の実行の間の条件関係は否定されるが、利益の帰属主体である Y は第 3 項 3 の論証と同様に、166 条 3 項の構成要件に該当すると解すべきであろう。それに対して、④事例では、外形的に Y が間接的に 166 条 3 項の構成要件を実

現したように見える。しかし、インサイダー取引の適用除外規定の趣旨に鑑みると、このような事例においても Y に刑事責任を帰属させたり、課徴金を課したりするのは、ポートフォリオを構築する際に AI を用いる利用者にとって過度な負担を強いるおそれがある。特に刑事事例では厳格な証明により、裁判官において違反行為に係る事実の存在につき合理的な疑いを容れない程度に確信を抱きうる程度の立証を要するのに対し、課徴金手続における審理では刑事訴訟のように合理的な疑いを超える立証までは要求されず、民事訴訟で要求されるような立証水準で足りる。さらに、この審理は刑事手続のような起訴便宜主義も存在しない。実務的見解でも「法令違反行為に対して審判手続という裁判に似た手続きを経て、行政処分として課徴金を課すものであり、刑事裁判に比べれば立証の程度が少なくて済む」[355]（証券取引等監視委員会）という。これらの状況を踏まえると、利用者にとっては、④事例においても課徴金対象となる可能性があり、事後的に取消訴訟を提起できるとはいえ、手続に係る負担、適用除外要件の抗弁を審理で行うのが被審人、すなわち利用者であることを考慮すれば、こうした状況は利用者にとって過大な負担となりうる。

そうした状況を未然に防ぐためには、検証可能な AI・アルゴリズムの構築の義務付けを製造者に課し、それを充足した上で、利用者が未公開重要事実を知ったことと取引実行条件関係が否定される場合を適用除外規定の中に創設すべきである。

第 4 目　小括

AI・アルゴリズムを利用したポートフォリオ構築において、これを介した金融商品取引がインサイダー取引を実現したとされる場合、そのポートフォリオ利用者の行為がインサイダー取引に該当しうるか否かは、現行金商法の解釈の範疇で解決可能である。具体的には、その利用者の行為自体はインサイダー取引（金商法 166 条 1 項・3 項）に該当する。しかし、重要事実を「知った」利用者のポートフォリオが、学習により最適化された AI により、その利用者の知らないところで有価証券等の取引を行った場合でも、課徴金による処分もしくは刑事責任の対象となる可能性がある。その際に適用除外要件を用いられるよう、AI・アルゴリズムを利用した取引の場合はその AI プロセスを検証可能なものにすることを製造者に義務付け、利用者が未公開重要事実を知ったことと取引実行の条件関係が否定される場合には適用除外とするといった類型を新たに設ける必要がある。

第３款　協調的行為

第 1 項　問題の所在

近時では、AI・アルゴリズムが競争事業者の価格の調査や自社商品・サービスの価格設定に活用されつつある。例えば、航空業界におけるレベニューマネジメントに基づいて行われるダイナミック・プライシングなど、競争事業者の価格を収集して，それに対応して自社の

[355] 証券取引等監視委員会「証券取引等監視委員会の活動状況(平成 20 年度)」（2009 年）65 頁参照。なお、この年度以降に公表された年次報告書にはこの記述が削除されている。

価格を設定したり，需要を予測した上で，販売量を最大化する価格を設定したりするためにアルゴリズムが用いられることがある[356]。その一方で、自己学習型 AI・アルゴリズムによる価格協調行為が、競争事業者らの認識のないまま、互いに競争的価格を上回ることによって実現される可能性も指摘されている[357]。現状の技術では実現可能性は低いものの、特定条件下での機械学習によって部分的な共謀を組織的に学習するという理由で価格上昇を観察した事例も報告されており[358]、この点において独占禁止法（以下、「独禁法」とする）89 条 1 項 1 号所定の不当な取引制限罪の成否が検討される。本項ではまず、独禁法下の不当な取引制限罪を確認しつつ、この問題についての検討を行う。

第 2 項 独占禁止法における不当な取引制限罪

第 1 目　独占禁止法の概要と規制対象・エンフォースメント

独占禁止法は、公正かつ自由な競争を促進するため、競争を制限する又は公正な競争を阻害する談合・カルテルなどの共同行為による不当な取引制限、他の事業者の事業活動を排除又は支配することによる私的独占、合併・事業譲渡・株式取得等による企業結合、そして再販売価格維持行為、不当廉売、抱き合わせ、優越的地位の濫用等による不公正な取引方法という主な 4 つの類型を規制する。このうち、私的独占と不当な取引制限については、排除措置命令（独禁法 7 条）・課徴金納付命令（独禁法 7 条の 2）が課されるのが原則であり、違反した場合の刑事罰も法定されている（独禁法 89 条 1 項 1 号・2 号）が[359]、不公正な取引方法については直接に処罰する規定は設けられていない。過去に処罰された事例は，いずれもの不当な取引制限に関するものであって、その典型といえるのは価格カルテルと入札談合である。

不当な取引制限に対する措置は大別して、①競争秩序の回復を行うための事業者に対する排除措置命令、②不当な取引制限等の対象となった取引の売上額につき、事業者に対し一定額の国庫への納付を命じる課徴金納付命令、③個人に対する刑事罰及び両罰規定を通じた法人に対する刑事罰である。

このうち、排除措置命令とは「事業者に対し、当該行為の差止め、事業の一部の譲渡その他これらの規定に違反する行為を排除するために必要な措置」（独禁法 7 条 1 項）であり、具体的には、違反行為の取りやめ、違反行為を取りやめていることを確認する取締役会決議、取引先への通知、従業員への周知徹底、将来同様の行為を行わない旨の不作為命令、独占禁止法遵守マニュアルの作成や研修の実施等に加え、他社との販売価格改定の情報交換禁止や談合に関与した常業担当者の社内配置転換が命じられることが挙げられる[360]。

---

356 デジタル市場における競争政策に関する研究会「アルゴリズム／AI と競争政策」（2021 年）10 頁。

357 デジタル市場における競争政策に関する研究会・前掲（注 356）25 頁。

358 *Calvano et al.*, Artificial Intelligence, Algorithmic Pricining, and Collusion, American Economic Review, 3267(2020), 110(10).

359 斉藤ほか・前掲（注 306）226 頁（中里）。

360 斉藤ほか・前掲（注 306）248 頁（中里）。

課徴金納付命令とは、1970 年代のオイルショック以降の石油価格高騰による闇カルテルの多発を背景に、それまでは排除措置命令を中心に対応していた運用に対して、カルテルによって得られた利益を放置することこそが社会的公正に反するのではないかという疑義が向けられたことに端を発する。そこで、1977 年にカルテル・談合を行った事業者に対して課徴金納付を命じるという新たな制度を創設することでこの問題の解決を図った。この課徴金納付命令は、「カルテルを抑止するための実効性ある手段として、カルテルによる不当利益を享受する者に対して、その利得をはく奪するために、一定の経済的負担を強制する行政上の権限」と説明され、法定された算定率に従って一律かつ画一的に賦課するものであり、公正取引委員会の裁量は認められない。なお、2005 年独禁法改正では事業者が自ら関与したカルテル・入札談合について，その違反内容を公正取引委員会に自主的に報告した場合，課徴金が減免される課徴金減免制度（リニエンシー制度）が導入されている。これは、事業者自らがその違反内容を報告し，更に資料を提出することにより，カルテル・入札談合の発見を容易化し，事件の真相解明を効率的かつ効果的に行うことにより，競争秩序を早期に回復することを目的としている。ただし、その効果は行政上の措置である課徴金の減額又は免除で反映されるにとどまり、犯則調査手続及び刑事告発との関係は法律上の定めはなく、当制度を利用しても刑事責任を負う可能性が残り、当制度のインセンティブが失われるのではないかという指摘があること[361]に留意しなければならない。

刑事罰については、私的独占又は不当な取引制限があった場合には、その行為者を罰するのみならず、 その法人に対しても 5 億円以下の罰金刑を科すことができると規定する（独禁法 95 条）。これらの違反行為については、公正取引委員会の専属告発とされ、事件を刑事訴追すべきか否かの第一次的判断権が公正取引委員会に与えられているので、これは独立行政委員会たる公正取引委員会の職権行使の独立性の徴表と考えられる[362]。刑事告発が検察官の訴訟条件となっており（独禁法 96 条）、刑事告発については、公正取引委員会に裁量権限があると解されている[363]。

第 2 目　不当な取引制限罪の構成要件における「共同して」の要件

独禁法 89 条 1 項 1 号にいう「不当な取引制限」とは、「事業者が、契約、協定その他何らの名義をもつてするかを問わず、他の事業者と共同して対価を決定し、維持し、若しくは引き上げ、又は数量、技術、製品、設備若しくは取引の相手方を制限する等相互にその事業活動を拘束し、又は遂行することにより、公共の利益に反して、一定の取引分野における競争を実質的に制限すること」（独禁法 2 条 6 項）と定義される。その解釈論として議論されてきたのは同規定のうち「共同して」、「相互に事業活動を拘束し、又は遂行する」という行為要件と「一定の取引分野における競争を実質的に制限する」という結果要件であるが、価格

361 斉藤ほか・前掲（注 306）254 頁（中里）。

362 斉藤ほか・前掲（注 306）253 頁（中里）。

363 東京高判平成 5 年 5 月 21 日 高刑集 46 巻 2 号 108 頁。

調整に供する AI・アルゴリズムを利用する競争事業者間の認識のないまま、互いに競争的価格を上回ることによって価格協調行為が実現される事例において問題となるのは、「共同して」の要件である。

「共同して」というためには、他の共同行為者との間に、「意思連絡」が必要であるというのが一般的見解であったとされる[364]。これまで実際に刑事罰が科されてきた事案は、いずれも共同行為者間に明示的な合意が認められる事案であるが、この関連ではいわゆる意識的並行行為と区別する必要がある。例えば、A 社がある商品について値上げを行い、別の B 社、C 社、D 社がそれぞれ独自の判断でこれに追随して値上げを決定したとしても、意思の連絡がないことから各事業者の単独判断に基づく通常の経済活動であり、「共同して」の要件を満たすことにはならない[365]。「共同して」の要件を満たす、すなわち意思の連絡があったとするには、価格カルテルについての協定が締結されたり、会合での入札談合の基本ルールの決定がなされたりする場合が考えられる。しかし、近時ではこれに限らず、少人数での会合、携帯電話、電子メールを活用するなど、事業者間の情報交換と隠蔽手段が巧妙となっていることがある。このような事情を踏まえて、東京高判平成 7 年 9 月 25 日判タ 906 号 136 頁（東芝ケミカル事件差戻審）では、「『意思の連絡』とは複数事業者間で相互に同内容又は同種の対価の引上げを実施することを認識ないし予測し、これと歩調をそろえる意思があることを意味し、一方の対価引上げを他方が単に認識、認容するのみでは足りないが、事業者間相互で拘束し合うことを明示して合意することまでは必要でなく、相互に他の事業者の対価の引上げ行為を認識して暗黙のうちに認容することで足りると解するのが相当である」。「もともと『不当な取引制限』とされるような合意については、これを外部に明らかになるような形で形成することは避けようとの配慮が働くのがむしろ通常であり外部的にも明らかな形による合意が認められなければならないと解すると、法の規制を容易に潜脱することを許す結果になるのは見易い道理であるからこのような解釈では実情に対応し得ないことは明らかである。したがって、対価引上げがなされるに至った前後の諸事情を勘案して事業者の認識及び意思がどのようなものであったかを検討し、事業者相互間に共同の認識認容があるかどうかを判断すべきである」として、黙示による意思連絡でも「共同して」の要件を満たしうるとし、この判例法理が現在に至るまで確立されている[366]。例えば、会合において、主催者が出席した各社の値上げ方針の賛否を確認した際、特に異論を述べずにそのまま退席し、会合後に確認のあった方針どおりに価格引き上げを行った事業者であっても当該要件を充足するとされる[367]。

---

364 白石忠志『独占禁止法〔第 2 版〕』（有斐閣、2009 年）。

365 斉藤ほか・前掲（注 306）232 頁（中里）。

366 斉藤ほか・前掲（注 306）232 頁（中里）、泉水文雄『独占禁止法』（有斐閣、2022 年）202 頁以下、金井貴嗣・川濵昇・泉水文雄編『独占禁止法（第 6 版）』（弘文堂、2021 年）48 頁以下。

367 斉藤ほか・前掲（注 306）232 頁（中里）。その詳細については、武田邦宣「不当な取引制限における意思の連絡要件」日本経済法学会年報 37 号（2016 年）19 頁以下を参照。

第 3 項 AI・アルゴリズムによる価格協調と不当な取引制限罪の成否

では、価格調整に供される AI・アルゴリズムによる価格協調行為が、競争事業者らの認識のないまま、互いに競争的価格を上回ることによって実現された場合、当該 AI・アルゴリズムを利用した競争事業者は「共同し」ていたといえるのか。確かに上述した判例法理によれば、明示的な意思連絡がなくとも「共同した」と認められる余地はあるが、事業者はもっぱら当該 AI・アルゴリズムを利用して価格調整を独自に行ったにすぎないので、事業者間においておよそ黙示による意思連絡を認定するに足りる事実はないといえる。仮に当該 AI・アルゴリズムを利用すること自体がこれを認定する事実だとするならば、当該 AI・アルゴリズムを利用した事業者はその認識・認容なく協調的行為を競争事業者間で行ったという奇妙な帰結に至ってしまう[368]上に、この状況下で排除措置命令、課徴金、刑罰といった制裁を受ける可能性が常にあるとするならば、価格調整に利用する AI・アルゴリズムの便益は失われてしまうといってもよい。いずれにしても、事業者に不当な取引制限罪（独禁法 89 条 1 項 1 号）を認定するのは不適切である。

第 4 項 海外の議論と将来的な規制

利用事業者がその動作の詳細を理解できないブラックボックス性を帯びる価格調整 AI・アルゴリズムの場合においても、これによって惹起された結果が利用事業者に帰せられるかについては海外でも議論されている。例えば、事業者がその従業員の行為に対して責任を負うためには、「関係する事業のパートナーまたは主要な管理者による行為や知識は必要なく、事業を代表して行動する権限を持つ人物の行動で十分である」[369]という欧州裁判所の判例に従い、AI・アルゴリズムを自社の従業員が協調的行為に従事した場合と同視して、AI・アルゴリズムによる行為も利用事業者の行為と捉え得るとする考え方[370]や、AI・アルゴリズムの開発や利活用への委縮効果に配慮し、注意義務や予見可能性の合理的基準に反する場合のみ責任を負うこととすべきとする立場[371]がある。なお日本においても、事業者責任を検討する際、不真正不作為犯構成をベースに「公序に整合した形でのみ AI を利用する義務が事業者には課されており、その認知の期待可能性・不作為の責任は課されて」おり、「競争

---

368 例外的な場合として、競争関係にある事業者間において各々の自己学習 AI・アルゴリズムを用いれば価格が同調することを相互に認識しながら当該 AI・アルゴリズムを利用した場合には、意思の連絡が認められる可能性があると考えられる（デジタル市場における競争政策に関する研究会・前掲（注 356）26 頁脚注 51 参照）が、非常に限定された条件下でのみ「共同して」の要件が認定されえないことが窺える。

369 ECJ, Musique Diffusion française and Others v Commission, Judgment of 07.06.83, Joined Cases 100/80 to 103/80, para. 97

370 See *De la Autorité councurrence, Budeskartellamt,* Algorithms and Competition, 2019, p.58.

371 *Janka/Uhsler,* Antitrust 4.0, European Competition Law Review 2018, pp. 112 et seq. (121); *Salaschek/Serafimova*, Preissetzungsalgorithmen im Lichte von Art. 101 AEUV, Wirtschaft und Wettbewerb 2018, pp. 8 et seq. (15 et seq.).

機能を侵害するおそれのあるアルゴリズムをもったソフトウェアにつき、その利用を決定する、もしくは継続利用する行為・意図を先行行為と捉え、弊害に対する因果関係を検討し、違法性を評価する」という見解[372]がみられる。

しかし、これら見解において価格調整をする AI・アルゴリズムの利用者たる事業者の義務を論じるには問題が残る。AI・アルゴリズムによる行為を利用事業者の行為と捉える見解では、仮に事業者側で AI・アルゴリズムが検証不可能な挙動をした場合にも事業者の行為と同視するのでは事業者の負担が過大なものになりうるし、注意や予見可能性の合理的基準に反する場合にのみ責任を負うとすべき見解でも、求められる注意内容が明確でなければならない。それは「公序」といった不明確な概念に基づくべきでない。自動運転車の事故事例における製造者の刑事製造物責任の負責を検討する際に重視したように、価格協調に係る不当な取引制限行為に対する刑事責任のレベルまで見据えて検討する際には、なお事業者に対する明文の義務付けが必要である。そこでは、AI・アルゴリズムが競争機能を侵害しないように事業者が監視する義務を何らかの形で設けるべきであるものと思われる。例えば競争の機能の侵害のおそれが想定されるアルゴリズムを持つソフトウェアの利用について、一定の回避措置を期待し、需要者に対して当該 AI・アルゴリズムの透明性・説明責任確保を果たしたか否かをつうじて、義務違反との関係を問うというスキームが考えられる[373]。

第 5 項　小括

AI・アルゴリズムを利用した価格調整によって、競争事業者らの認識のないまま互いに競争的価格を上回ることが実現された場合、現行独禁法では不当な取引制限罪が成立する余地があるように見える。特にこの類型における「共同して」要件の解釈においては、外形的事実に基づいて認定されうるところにその問題があるが、少なくとも当該 AI・アルゴリズムを利用したという事実のみでこの要件を充足させるべきではない。とはいえ、一定条件の下では実際にこのようなことが起こりうることも実証されている以上は、競争法をつうじた規制が求められることは否定できない。しかし、不当な取引制限行為には排除措置命令のみならず、課徴金、さらには刑罰という法効果をもたらしうる構造であることを考慮すれば、当該 AI・アルゴリズムを利用する事業者に向けた明文の義務付けが必要であることも否定できないだろう。

第 3 節　コンピュータ領域の犯罪―行為客体としての AI

前節では AI 製品が刑法上の犯罪ではなく、特別刑法に属する証券犯罪や競争法違反に係る犯罪の構成要件を実現した場合の検討を行った。本節では、第 2 章や前節で取り扱ったよ

---

372 市川芳治「人工知能（AI）時代の競争法に関する一試論 ～“アルゴリズム”によるカルテル：欧米の最新事例からの示唆を受けて～（下）」国際商事法務 45 巻 2 号（2017 年）169 頁。

373 市川・前掲（注 372）168 頁参照。

うなAI製品が利用者の知らないところで各構成要件を実現した場合とは異なり、これまで日本の先行研究ではあまり意識されなかった、ネットワーク化されたAI製品が攻撃を受けた際の刑法上の評価について検討したい。

AIを搭載した製品に対し、その機能を害する行為が行われた場合、充足が考えられる刑法上の構成要件として、ドイツ刑法では、データ探知罪（202条a)、データ取得罪（202条b)、データ探知罪及びデータ探知罪、データ変更罪（303条a)、コンピュータ破壊罪（303条b)、ないしはコンピュータ詐欺罪（263条a）が挙げられる[374]。日本におけるコンピュータ犯罪の類型とは異なるものがあるが、そこからの示唆は日本刑法においても参考となるものがあろう。以下では、日本刑法における、いわゆるコンピュータ犯罪の制定経緯を踏まえ[375]、その上で、ドイツ刑法における犯罪類型と重なり合う、不正アクセス罪・通信の秘密侵害罪、電子計算機業務妨害罪、電子計算機使用等詐欺罪の適用可否を検討する。

第1款　コンピュータ刑法の制定経緯

コンピュータの普及、それに伴う情報処理の高度化の進展、情報通信技術の進展によりインターネットがグローバル単位で形成され、今やインターネットに接続されたコンピュータによる情報通信は社会における不可欠なインフラストラクチャーとして機能している。このような社会において、コンピュータやインターネットを悪用した犯罪が懸念され、1987年の刑法一部改正においては、情報処理組織において用いられる電磁的記録について、その不正作出および供用ならびに毀棄を処罰する規定を電磁的記録不正作出等罪（刑法161条の2)、電子情報処理組織による大量かつ迅速な大量迅速な情報により行われる業務を妨害する行為を処罰する規定を電子計算機業務妨害罪（刑法234条の2)、債権、債務の決済等か電磁的記録を用いて自動的に行われる事務処理の形態を利用して財産上不法の利益を得る行為を処罰する規定を電子計算機使用等詐欺罪（刑法246条の2)、及び電磁的記録毀棄罪（刑法258条・259条）が創設された。

1987年改正で残された議論は、いわゆるハッカーを含むコンビュータシステムに対する不正アクセスという類型を考える必要性の有無であった[376]。しかしながら、不正アクセスは上記4つの類型の予備的手段であり、コンピュータの情報処理機能に対する実質的加害とは必ずしもいえず、むしろ、立法的対応を考える上で実質的加害を個々にとらえて新たな構成要件を創設すべきか、もしくは実質的加害の手段となるべき行為をとらえて構成要件とするのが適当かという問題と位置付けることが相当であるように思われたため[377]、当該不正アクセス行為に対応する構成要件の創設は見送られた。

---

374 *Günther*, a.a.O. (fn.176), S.228.

375 安富潔「情報化社会における刑事立法の役割―コンピュータ犯罪からサイバー犯罪へ―」慶応法学42号（2019年）379頁以下も参照。

376 米澤慶治『刑法一部改正法の解説』（立花書房、1989年）12頁。

377 米澤・前掲（注376）12頁。

その後、1997 年 6 月 22 日に米国で開催された 8 ヶ国デンヴァー・サミットにおける「コミュニケ」では、1996 年のリヨン・サミットにおける「国際組織犯罪に関する 40 の勧告（いわゆるリヨン・グループ 40 の勧告）」を受け、その 40 において、「1 つは、コンピュータ及び電気通信技術に対して国境を越えて介入するようなハイテク犯罪者についての捜査、訴追及び処罰」を、「もう 1 つは、犯罪者の所在地にかかわらず、すべての政府がハイテク犯罪に対応する技術的及び法的能力を有することとなる体制」を構築することを採択した[378]。さらに、同年 12 月 10 日に米国ワシントン DC で開催された G8 司法・内務閣僚級会合では、「ハイテク犯罪と闘うための原則と行動計画」が採択され、例えば、「電気通信及びコンピュータ・システムの濫用を適切に犯罪化しハイテク犯罪の捜査を促進することを確保するため、我々の法制度を見直す」[379]などといった行動計画が策定された。

このような国際情勢も相俟って、諸外国では不正アクセス行為を含めたコンピュータ犯罪、ないしはハイテク犯罪に関する法整備が進められていった[380]が、日本においては先述の通り、不正アクセス行為の処罰規定は存在していなかった。とりわけ、当該規定の不存在にため、そして国際捜査共助が双方の可罰性を要件として実施されていることにより、仮に不正アクセスが日本を経由してなされた場合、関係国の捜査に協力できないため、日本が国際的なハイテク犯罪対策おける抜け道となる事態が生じてしまうという状況にあったということが不正アクセス罪の創設の理由となった[381]。そこで、これらの事情をもとに、ハイテク犯罪を防止するとともに電気通信に関する秩序を維持し、もって高度情報通信社会の健全な発展に寄与することを目的として不正アクセス禁止法が制定されるに至ったのである。

21 世紀に入り、2011 年には、いわゆるサイバー犯罪その他の情報処理の高度化に伴う犯罪及び強制執行を妨害する犯罪の実情に鑑みて、情報処理の高度化に伴う犯罪に適切に対処するとともに欧州評議会の「サイバー犯罪に関する条約」（以下「サイバー犯罪条約」とする）を締結するため，罰則及び手続法の整備を行うほか，強制執行を妨害する犯罪に適切に対処するため，罰則の整備を行う必要性をもとに、「情報処理の高度化等に対処するための刑法等の一部を改正する法律」（平成 23 年法律第 7 号）が成立した[382]。それにより、刑法では主に、不正指令電磁的記録に関する罪の新設（刑法 168 条の 2・3)、わいせつ物頒布等の罪の構成要件の拡充等(刑法 175 条)、電子計算機損壊等業務妨害罪の未遂犯処罰規定の新設（刑法 234 条の 2 第 2 項）が、刑事訴訟法上では主に、電気通信回線で接続している記録媒体からの複写の導入（刑事訴訟法 99 条 2 項,218 条 2 項)、記録命令付差押えの新設(刑事訴訟法 99 条の 2、218 条 1 項)、電磁的記録に係る記録媒体の差押えの執行方法の整備(刑事

---

378 外務省「8 ヵ国デンヴァー・サミット コミュニケ（仮訳)」(1997 年)。

379 外務省「8 カ国司法・内務閣僚級会合 1997 年 12 月 9-10 日　コミュニケ（仮訳)」(1998 年)。

380 安富・前掲（注 375）389 頁。

381 安富・前掲（注 375）390 頁脚注 25 の指摘である。

382 杉山徳明・吉田雅之「『情報処理の高度化等に対処するための刑法等の一部を改正する法律』について」曹時 64 巻 4 号 2 頁など。

訴訟法 110 条の 2、222 条 1 項)、保全要請に関する規定の整備（刑事訴訟法 197 条 3 項~5 項）が改正されている。

このように、「コンピュータ刑法」に位置付けられる各類型の制定過程はこの 30 年間に行われきたものである。いずれも、コンピュータの普及、及びそれに伴う情報処理の高度化による新たな犯罪に対応するものであった。以下では、それらの犯罪類型のうち、AI を搭載した機械に対する「コンピュータ刑法」において問題となりうる類型を、Günther によるドイツ刑法における検討と比較しながら、その要件解釈を基にしながら検討する。その行為類型としては以下のものが考えられる。すなわち、ある AI を搭載した機械をハッキングする行為、その内部データを変更または破壊する行為、そしていわゆるインテリジェント・エージェントを利用したコンピュータ詐欺行為である。

第 2 款 AI 製品に対するデータ探知・取得と不正アクセス

本項においては、ある者がネットワーク化（IoT 化）された AI を搭載した製品（以下、「当該 AI 製品」とする）にハッキングし、学習によりその利用者に最適化された内部データを探知もしくは取得したという事例をベースに刑法上の検討を行う[383]。

第 1 項　ドイツ刑法下の検討

無権限アクセスをした場合、ドイツ刑法では 202 条 a（データ探知[Ausspähen von Daten]）の構成要件に該当しうる[384]。この構成要件の保護法益は、データ処分権限者の秘密保持利益とされる[385]。また刑法 202 条 a 第 2 項では、データの定義規定が設けられており、データは「電子的、磁気的、またはその他の直接的に知覚不可能な形で保存または送信されるもの」とされる。

AI 製品を利用する際には、制御ソフトウェアにおいて、システム自身によって収集したユーザーの情報もしくは、ユーザーのコードナンバーなど、あらゆる点で情報が常に処理されており、実装されたソフトウェアでさえも処理の手段となることもある[386]。また、新たな種類のデータ保存も構成要件に該当しうるので[387]、行為客体となるのは、必ずしも市販のコンピュータ、もしくはハードディスク、USB メモリ、DVD などの古典的なデータ記憶媒体に保存されたデータに限らず、原則として保存または送信されているすべてのデータが行為客体とされる[388]。ただし、そのデータがある利用者のために予定されたものや、アクセスに対して特別に保護されたものであることは要求されておらず、むしろ「非決定性」は、処

---

383 本項では、問題となる AI 製品が学習機能を有するか否かは問題にしない。

384 法定刑は 3 年以下の自由刑または罰金刑である（§ 202 a (1) StGB）。

385 *Leupold/Glossner* (Hrsg.), Münchener Anwaltshandbuch IT-Recht, C.H.Beck 2011, Teil 10, Rn. 74.

386 *Günther*, a.a.O. (fn.176), S.229.

387 *Hilgendorf/Valerius*, Computer- und lnternetstrafrecht, Rn. 539.

388 *Günther*, a.a.O. (fn.176), S.229

分権限者の意思に従って評価される[389]。そこでは、構成要件を限定する観点から[390]、不正アクセスに対する特別な保護というメルクマールが尊重される。このようなセキュリティは、ハードウェアやプログラムで機械的に実装することも可能であり[391]、特に AI・ロボット工学の分野では、将来的にハードウェア側とソフトウェア側のセキュリティがシステムに統合されることが予想される。とはいえ、現在の AI 開発・研究の段階では、このようなセキュリティはむしろ非典型的なものと思われるが、それらの利用が進むにつれて、システムで処理されたデータに対する権限者の秘密保持利益も高まることが考えられる。例えば健康、嗜好に関するデータ、もしくは位置情報データなど、AI 製品の利用者やその特性に関する情報を含むデータである場合にはさらなる保護が望まれるといってもよい[392]。2007 年のコンピュータ犯罪撲滅のための第 41 次刑法改正では、刑法 202 条 a の処罰範囲が拡大され、データの取得だけでなく、その予備行為であるデータへのアクセス取得、すなわちシステムへの侵入も構成要件として追加された[393]。例えば、データの純粋な閲覧だけでも、データ記憶媒体のコピー、マルウェアのインストール、また、暗号化されているが利用可能でないデータへの純粋なアクセスによって入手できるにすぎない状況においても構成要件は充足しうる[394]。いずれにしても、冒頭の事例のように、ある者が AI を搭載した機械へハッキングを行い、それによりデータへのアクセスが可能となった場合、刑法 202 条 a の構成要件に該当しうることになる[395]。

データ取得行為については、伝達されたもしくは電磁的に発せられたデータの取得を処罰するドイツ刑法 202 条 b の構成要件に該当しうる[396]。この規定は受け皿構成要件で、処分権限者の秘密保持利益を保護するものとされる[397]。そこで、このデータがどのように伝送されるか、有線か無線か[398]、あるいは、パブリックネットワークかプライベートネットワークかは重要ではない[399]。例えば、取得されたステータスメッセージ、製造者、契約相手、もしくはユーザーに送信された AI 製品の照会は行為客体に適するだろう。また、電磁的放射も刑法第 202 条 b の対象となり、AI 製品の電磁的放射の場合に測定される「サイドチャネル攻撃」も本罪の構成要件に該当しうる[400]。 もちろん、刑法第 202 条 b は、AI 製品からパ

---

389 *Fischer*, a.a.O. (fn.181), § 202a, Rn. 7 f.これは、データが防護されていたり、行為者にとって予定されていたりと、データの性質は必ずしも決まっているわけではないということで「非決定性」という。

390 *Fischer*, a.a.O. (fn.181), § 202a. Rn. 8; *Hilgendorf/Valerius*, a.a.O. (fn.387), Rn. 546 ff.

391 *Fischer*, a.a.O. (fn.181), § 202a. Rn. 9.

392 *Günther*, a.a.O. (fn.176), S.241.

393 *Ernst*, Das neue Computersstrafrecht, NJW 2007, 2661 f.

394 Vgl. *Fischer*, a.a.O. (fn.181), § 202a. Rn. 10 ff

395 *Günther*, a.a.O. (fn.176), S.229. なお、ドイツ刑法 303 条には器物損壊罪が規定されている。

396 *Leupold/Glossner*, a.a.O. (fn.385), Teil 10. Rn. 89 f.

397 *Hilgendorf*, Das neue Computerstrafrecht, in: *Hilgendorf* (Hrsg.), Dimensionen des IT-Rechts (Das Strafrecht vor neuen Herausforderungen). Logos 2008, S. 5.

398 例えば W-LAN、Bluetooth、RFID、赤外線などが考えられる。Vgl. *Günther*, a.a.O. (fn.176), S.240.

399 *Fischer*, a.a.O. (fn.181), § 202b, Rn. 3.

400 *Hilgendorf*, a.a.O. (fn.397), S. 6.

ートナー・ステーションへのデータ伝送だけでなく、パートナー・ステーションからロボットへのデータ伝送もカバーする。このデータは、保存したり、認識したりすることによって入手されるものであり[401]、ハッキングによってアクセスできる可能性のみでは可罰性はない[402]。したがって、AI 製品に送信されたデータもしくは AI 製品から発出するデータが伝送中に取得された限りで、刑法第 202 条 b の構成要件に該当する。

第 2 項　日本法における適用

日本法下においてこの事例が関連するのは、アクセス行為につき不正アクセス罪（不正アクセス禁止法 3 条・11 条）が、データ取得行為につき電気通信事業法 179 条 1 項または有線電気通信法 9 条が関連する。

第 1 目　アクセス行為と不正アクセス罪

不正アクセス罪とは不正アクセス禁止法 2 条 4 項によると、「アクセス制御機能を有する特定電子計算機に電気通信回線を通じて当該アクセス制御機能に係る他人の識別符号を入力して当該特定電子計算機を作動させ、当該アクセス制御機能により制限されている特定利用をし得る状態にさせる行為」（1 号）という識別符号を盗用する類型、「アクセス制御機能を有する特定電子計算機に電気通信回線を通じて当該アクセス制御機能による特定利用の制限を免れることができる情報又は指令を入力して当該特定電子計算機を作動させ、その制限されている特定利用をし得る状態にさせる行為」（2 号）、「電気通信回線を介して接続された他の特定電子計算機が有するアクセス制御機能によりその特定利用を制限されている特定電子計算機に電気通信回線を通じてその制限を免れることができる情報又は指令を入力して当該特定電子計算機を作動させ、その制限されている特定利用をし得る状態にさせる行為」（3 号）というセキュリティ・ホールを攻撃する類型が規定されている。この犯罪類型の目的は、ID・パスワードやバイオメトリクス情報といった識別符号から構成される認証システムの保護にあるとされ、そのような認証システムのことを アクセス制御機能と呼び、アクセス制御機能に対する社会的信頼が保護法益とされる[403]。ここで注意しなければならないのは、アクセス制御機能を特定できない場合やそのアクセス制御機能の防御効果を無効化・迂回しないセキュリティ・ホール攻撃は、不正アクセス禁止法が保護しようとしている利用者識別とは関係のないサイバー攻撃ということになり、不正アクセス罪では処罰されない行為であるということである[404]。言い換えれば、不正アクセス罪はサイバー攻

---

401 *Leupold/Glossner*, a.a.O. (fn.385), Teil 10, Rn. 91.

402 *Fischer*, a.a.O. (fn.181), § 202b, Rn. 5; *Hilgendorf/Valerius*, a.a.O. (fn.387), Rn. 571.

403 鎮目征樹・西貝吉晃・北條孝佳『情報刑法Ⅰ　サイバーセキュリティ関連犯罪』（弘文堂、2022 年）143 頁（西貝）。

404 鎮目ほか・前掲（注 403）143 頁（西貝）。他に、スタンドアロンのコンピュータ（ネットワークに接続されていないコンピュータ）を無断で使用する行為や、ネットワークに接続されアクセス制御機能により特定利用が制限されているコンピュータであっても当該コンピュータのキーボード（コンソール）を直

撃全般を把握しているわけではない。

それでは冒頭で挙げた行為類型のように、AI を搭載した製品に第三者がハッキングを行い、当該製品の利用者に最適化されたデータにアクセス、取得した場合の刑法上の評価はどのようになるか。不正アクセス罪が対象とするのはアクセス行為のみであるため、当該アクセス行為が不正アクセス禁止法 3 条所定の不正アクセスに該当するかが問題である。ここでは、当該 AI 製品内の利用者に最適化された内部データにアクセス制御機能が施されていれば、この要件は充足しうる。取得行為自体は次で取り扱う通信の秘密侵害罪の内容となる。

第 2 目　データ取得と電気通信の秘密侵害罪

電気通信の秘密を侵害する罪として冒頭事例に関連するのは、電気通信事業法 179 条、有線電気通信法 9 条（罰則：14 条）である。

電気通信事業法 179 条は、電気通信事業者の取扱中に係る通信の秘密を侵した者に対し 2 年以下の懲役刑または 100 万円以下の罰金刑（1 項）、それが電気通信事業に従事する者である場合には 3 年以下の懲役刑または 200 万円以下の罰金刑（2 項）、及び未遂処罰規定（3 項）をそれぞれ規定する。また、有線電気通信法 9 条は「有線電気通信の秘密は、侵してはならない」と規定して、同法 14 条では、有線電気通信の秘密を侵した者には 2 年以下の懲役刑または 50 万円以下の罰金刑（1 項）、それが有線電気通信の業務従事者の場合では 3 年以下の懲役刑または 100 万円以下の罰金刑（2 項）、未遂処罰規定（3 項）、及び国外犯処罰規定（4 項）が規定される。同罪の保護法益は、①私人のプライバシーの保護と思想、表現の自由の保障、②それに伴う電気通信業務の円滑適正な運用とそれに対する国民の信頼の確保を図ることを内容とする[405]。また、有線電気通信法 14 条 1 項所定の罪については、業務関連性を有しない通信の保護をも目的にするため保護法益は前記①に限定される。

電気通信事業法 179 条の「電気通信事業者の取扱中に係る」の定義[406]として、「電気通信事業者の取扱い中」とは発信者が通信を発した時点から受信者がその通信を受ける時点までの間をいい、電気通信事業者の管理支配下にある状態のものを指す。また「取扱中」に関わる通信の範囲とは、伝達行為が終了した後の情報も保護の対象となり、通信終了後にも電気通信事業者が保管している通信内容に関する記録、通信記録なども保護の対象になるとされている。原則、内的なコミュニケーションである通信については、本人の意思を尊重すべく、電話や電子メールなどの特定者間の通信は秘密性が推定される。これに対して、電子掲示板やホームページに掲載された情報など、不特定の者に対して表示することを目的と

接操作して無断で使用する行為も本罪の対象外である（警察庁サイバー犯罪対策プロジェクト「不正アクセス行為の禁止等に関する法律の解説」（https://www.npa.go.jp/cyber/legislation/pdf/1_kAIsetsu.pdf, 最終アクセス 2022 年 11 月 28 日）7 頁参照）。

405 角田正紀「判批」判評 356 号（1988 年）77 -78 頁、伊藤榮樹・小野慶二・荘司邦雄『注釈特別刑法 第 6 巻 交通法・通信法編 Ⅱ』（立花書房、1994 年）375 頁（河上）。

406 以下の説明は、多賀谷一照監修 電気通信事業法研究会編著『電気通信事業法逐条解説 改訂版』（2019 年、一般財団法人情報通信振興会）34 頁以下を参照。

した通信の内容は対象外とされる。「通信の秘密」の範囲は、通信内容のみならず、通信の日時、場所、回数、当事者の氏名、住所、電話番号など、通信の意味内容が推し量れるような事項すべてを含むと考えられ、「通信の秘密を侵す」とは、①通信当事者以外の第三者が積極的意思を持って知得しようとすること、②第三者にとどまっている秘密をその者が漏洩すること、及び窃用することも、それぞれ独立して秘密を侵すことに該当するとされている[407]。その具体例として、データの取得に相当する行為（知得）が捕捉されることには争いはないが、ここでいう知得とは、故意にかつ積極的に知得することであるから、偶然に目にふれ、または適法に知得することだけでは秘密の侵害という概念には含まれない。ここで取得（知得）というときに、取得したデータの意味内容を知る必要はないとされる[408]。つまり、通信の秘密の知得というためにはデータを取得することが必要なのである。

冒頭事例におけるデータ取得行為は当該 AI 製品の通信形態によって異なるが、いずれにしてもその利用者に最適化されたデータを取得する行為はその後の取得データの利用形態によらず上記類型（電気通信事業法 179 条、有線電気通信法 14 条）に該当するものと考えられる。その際、対象となるデータ取得が「秘密」の侵害たる「秘密」に該当するか否かについては、通信の秘密が、通信の内容だけでなく、例えば発信者の氏名・住所、発信回数・日時等通信の外形情報・通信の存在に関する事項も保護の対象とされるため[409]、ドイツ刑法でいうデータ取得のデータ概念のように処分権限を有する者の意思に従って評価されるわけではない。日本法におけるデータ取得行為自体は、インターネットに接続されている製品に学習能力を有する AI が搭載されているか否かによらず、上記類型に該当する。

まとめると、冒頭事例における行為者は当該 AI 製品の内部データにアクセスする行為につき不正アクセス罪が、その取得行為につき通信の秘密侵害罪が成立することとなり、両罪は併合罪の関係に立つ[410]。

---

407 葛西まゆこ「イントロダクション 憲法学から見た通信の秘密」警察学論集 66 巻 2 号（2013 年）130 頁以下。

408 鎮目ほか・前掲（注 403）189 頁以下（西貝）によると、「〈暗号化されたデータを取得したが、それの復号ができない場合であっても通信の秘密侵害になる〉という理解」も存在することから、これは「通信がデータのやり取りを意味し、データのやり取り自体が保護されていることと整合的な理解だといえる」という。

409 鎮目ほか・前掲（注 403）180 頁（西貝）。

410 最決平成 19 年 8 月 8 日刑集 61 巻 5 号 576 頁参照。さらに、前田巌「不正アクセス行為の禁止等に関する法律 8 条 1 号の罪と私電磁的記録不正作出罪との罪数関係」曹時 62 巻 10 号 165 頁によると、「不正アクセス行為には，それ自体を目的として行われるハッキングや処罰規定のない電子情報へのアクセスそのものを目的とする場合等を類型的に予定することができるのであり，不正アクセス行為を手段として電磁的記録不正作出・同供用や電子計算機使用詐欺、電子計算機損壊等業務妨害等に及ぶことが類型的に予定されているとまではいえない」という。

第 3 款 AI 製品に対するデータ変更・コンピュータ破壊

本項では、学習能力を有する AI を搭載した製品に対し、あるハッカーがその AI 製品にハッキングを行い、その内部データを変更したり、消去したりする事例における刑法上の検討を行う。

第 1 項　ドイツ刑法における議論

第 1 目　データ変更罪

刑法 303 条 a によると、データを違法に消去、隠匿、使用不能にする、または変更することが刑罰の対象とする。当該規範の体系的位置や条文の文言に基づくと、器物損壊と類似する[411]。本罪の保護法益は、データストレージに含まれる情報に対する権限者の処分権である[412]。本条における「データ」概念は、刑法 202 条 a 第 2 項で規定されるように「他人」という書かれざる構成要件要素によって規定される[413]。このデータの「他人性」とは、データに関する使用、処理、または削除する他者の権利がある場合のことをいうが[414]、AI の領域では、メールのフィルタリング行為がデータ変更に該当するか否かという、いわゆるインターネット上のメールボックス問題での議論に倣い、AI のネットワーク化における問題が生じうる。ここではさらに、データの伝送における処分権の有無や、いつそれが確立されたのかを明確にすべきであり、それゆえデータがアドレスされ、AI システムによって呼び出すことができる場合には、すでに処分権があると考えるべきかという問題が生じるが、少なくともデータ提供前の事前のフィルタリングや処理は、刑法 303 条 a に基づくデータ変更には該当しないと考えられる[415]。

構成要件的行為として、本条では、データの削除、隠匿、使用不能にすること、及び改変を認めている。その種類は、刑法 303 条による器物損壊に依拠するもので、内容上一部重なり合うところがある[416]。データの削除とは、完全かつ再現できない程度に識別できなくすることであり[417]、データの隠匿とは、データがその権限者から隔離され、権限者による利用が取るに足らないと言えない期間である場合に、データを利用不可能にすることをいう[418]。またデータの使用不能とは、データが通常に即して利用し得ない場合に、利用可能性が制限されていることをいい[419]、データの改変は、改変前のデータと異なる状態がもたらされたことをいう[420]。また予備段階についても、刑法 303 条 a 第 3 項により処罰の対象となる。

---

411 *Günther*, a.a.O.(fn.176), S.231
412 *Fischer*, a.a.O. (fn.181), § 303a, Rn. 2. m.w.N.
413 MüKo-StGB/*Wieck-Noodt*, § 303a StGB. Rn. 9.
414 *Hilgendorf/Valerius*, a.a.O. (fn.387), Rn. 588, *Fischer*, a.a.O. (fn.181), § 303a, Rn. 4 f.; MüKo-StGB/*Wieck-Noodt*, § 303a StGB, Rn. 9 f.
415 Vgl. *Fischer*, a.a.O. (fn.181), § 303a. Rn. 7
416 *Fischer*, a.a.O. (fn.181), § 303a. Rn. 8 ff.
417 MüKo-StGB/*Wieck-Noodt.* § 303a StGB. Rn. 12, m.w.N.
418 *Fischer,* a.a.O. (fn.181), § 303a, Rn. 10
419 BT-Drs. 10/4728, S. 36; MüKo-StGB/*Wieck-Noodt*, § 303a StGB, Rn. 9 f. m.w.N.
420 BT-Drs. 10/4728. S. 36.

刑法 303 条 a を適用しうる状況は、例えば無許可の第三者による AI 製品のデータストレージへのアクセスである。データに対して無権限の者が AI 製品をハッキングした後（ドイツ刑法 202 条 a)、上記のような構成要件的行為を遂行した場合、ドイツ刑法 303 条 a の構成要件に該当しうる。ただし、データの削除・隠匿・使用不能化もしくは改変が無権限者の無権限アクセスに伴う行為結果なのか、それとも AI の学習によるものなのかが不明確な場合が考えられる。このように因果関係が不明確な場合、構成要件的結果が発生しているにもかかわらず、その無権限者に対してはドイツ刑法 303 条 a の既遂罪の適用が否定され未遂罪（ドイツ刑法 303 条 a 第 3 項）の適用にとどまる。そうすると、当該 AI 製品にハッキングを行った無権限者が AI の学習を引き合いに出して既遂罪の適用を免れようとすることが考えられるだろう。そこで重要となるのが**説明可能な AI** の構想であり、事例におけるデータ変更等が AI の学習によってもたらされたか否かを証明できるように開発を行うことが、この類型の保護法益であるデータ権限者の処分権を保障することになる。もし AI の学習ではなく無権限者によってデータ変更がもたらされたといえるならば、刑法 303 条 a と刑法 202 条 a が成立し、両罪は観念的競合となる[421]。

第 2 目　コンピュータ破壊罪

コンピュータ・システムが破壊されたときに必ず考慮されるのが、刑法 303 条 b 所定のコンピュータ破壊である。本罪は、データ処理の適切な機能化に対する運用者または利用者の利益を保護するものとされる[422]。

刑法 303 条 b 第 1 項によると、データ処理が破壊されなければならない。このデータ処理とは、技術的に理解されるべきものであり、狭義の伝送、入力、処理など、電子計算処理のあらゆる形態を含むものとされる[423]。データ処理は個別のデータ処理経過のみを記述するという限界づけはなされないが[424]、AI 製品は非常に複雑かつ高度に自動化されたシステムを有することが多いので、通常では個別のプロセスのみが妨害されることはなく、「あるデータ処理」が妨害されることになる。ここで重要となるのは技術的見地であるため、例えば演算処理が内部でのみ実行され、限定された範囲にしか現れず、そのようなシステムが典型的に従来的装置の外観を呈しているにすぎない場合であっても、その経過はデータ処理に該当する[425]。データ処理装置とは処理が行われる機能単位のことを意味し、そこには AI 製品も含まれるとしてもよい[426]。さらにデータ処理経過は、他者にとって本質的に重要でなければならない。2007 年改正以降では私的なデータ処理も刑法第 303 条 b によって保護さ

421 *Fischer,* a.a.O. (fn.181), § 303a, Rn. 18.
422 保護法益の変化については BT-Drs. 16/3656, S. 13.
423 *Fischer,* a.a.O. (fn.181). § 303a, Rn. 4 f.
424 たとえば、*Lackner/Kühl*, Strafgesetzbuch: StGB, 29. Aufl.,C.H.Beck 2018, § 303b. Rn. 2; *Fischer,* a.a.O. (fn.181), § 303a. Rn. 4..
425 *Günther*, a.a.O. (fn.176), S.232.
426 *Günther*, a.a.O. (fn.176), S.233.

れているが、その本質的な意義はそれぞれのタスクまたは組織がデータ処理の性能に全面的または少なくとも大部分を依存している場合にある[427]。具体的には、基本的にデータ処理経過の観点から判断され、大別して企業や行政機関における場合と私人における場合とで区別されている[428]。前者では、例えば、人事管理、生産管理、購入や売却、物流管理、会計、会社の計画及び戦略決定の管理に使われるデータ処理、さらには E メールやウェブサイト上でのメッセージのやりとりも重要な意味をもつとされる[429]。後者は、私人の生活形成にとって中心的な機能を示しているか否かが重要であるとされる[430]。そのため、本質性要件は具体的状況における AI 製品の用途に応じて決定される必要がある。例えば介護や警備のための AI 製品であれば認められるかもしれないが、純粋に玩具としての役割を果たすにすぎない AI 製品には存在しない。付言すると、デュアル・ユースの AI 製品の場合はその線引きが困難であり、本質性要件を充足するか否か疑わしい場合には、利用時の重点が決め手となる。少なくとも、個人の生活を日常的に支援する AI 搭載の家庭用ロボットや介護に供する介護ロボットの場合は、本質性が認められてもよいと思われる[431]。

構成要件的行為は、刑法 303 条 a によるデータ改変（刑法 303 条 b 第 1 項 1 号)、不利益をもたらす目的でのデータの入力と伝達（刑法 303 条 b 第 1 項 2 号)、データ処理設備またはデータ媒体の破壊、損壊、使用不能化、除去または変更（刑法 303 条 b 第 1 項 3 号）である。2007 年以降、刑法第 303 条 b 第 5 項は、刑法第 202 条 c を参照して、コンピュータ破壊の予備も刑罰の対象としている。コンピュータ破壊は、前述のような構成要件的行為が AI 製品に対して遂行された場合に適用される。

第 2 項　日本刑法における検討

第 1 目　器物損壊罪の適用可否

想定する事例に関し、日本刑法においては、ドイツ刑法 303 条 a や同条 b のようなデータ変更やコンピュータ破壊のような構成要件は存在しないが、器物損壊罪（刑法 261 条）の適用の余地はあった。その先例としては、東京地裁平成 23 年 7 月 20 日 判タ 1393 号 366 頁（イカタコウイルス事件）が挙げられる[432]。

その判示において、「損壊」概念については、最判昭和 25 年 4 月 21 日 刑集 4 巻 4 号 655

---

427 *Fischer*, a.a.O. (fn.181), Rn. 6

428 SK-StGB, 9. Aufl. 2017, §303b Rn. 11. Spannbrucker, 75.

429 *Lackner/Kühl*, a.a.O. (fn.424), §303b Rn. 10

430 西貝吉晃「コンピュータ・サボタージュ罪 刑法 303 条 b」刑事法ジャーナル 71 号（2022 年）100 頁。

431 *Günther*, a.a.O. (fn.176), S.233.

432 本判例の評釈として、園田寿「『イカタコ事件』について」：器物損壊罪における「損壊」の概念〈判例批評〉甲南法務研究 8 号 103 頁、浅田和茂「ファイル共有ソフト利用者に「イカタコウィルス」を受信・実行させた行為が器物損壊罪に当たるとされた事例」新・判例解説 Watch（法学セミナー増刊）11 号 135 頁、森住信人「イカ／タコウィルス事件：ソフトウェアの改変と器物損壊罪の成否〈刑事裁判例批評 268〉」刑事法ジャーナル 41 号 211 頁がある。

頁を参照しつつ「物質的に物の全部又は一部を害し，あるいは物の本来の効用を失わせる行為をいう…。すなわち，器物損壊には，物自体を物理的に破壊する態様と物が持つ効用を侵害する態様があるが，後者の場合，『損壊』が成立するかどうかは，客体の効用を可罰的な程度に侵害したかどうかによって判断すべきであり，その効用侵害が一時的なものではないか，原状回復の難易をも考慮して検討すべきである」とし、「原状回復の容易性について判断する場合、利用者のコンピュータに関する知識レベルは様々であるところ、「損壊の成否は飽くまで社会通念に照らして判断すべきであるから、その難易は、パソコンの一般的な利用者を基準に判断すべきである」という規範を示し、その上で裁判所は、ハードディスクには保存されているデータを随時読み出せる機能（読み出し機能）と新たにデータを何度でも書き込める機能（書き込み機能）があることを定義し、「本件ウィルスにより，各被害者のハードディスクは，使用不能となったファイルが 保存されていた部分について読み出し機能が害された」「本件ウィルスの実行状態を止めない限り，ファイルを書き込んで保存しておくことは事実上不可能であり，ハードディスクの書き込み機能は害された」と判示した。

しかし、本件事案においては疑問が呈されることが多い[433]。というのも、器物損壊罪（261条）の客体は、他人の「物」であり、公用文書等毀棄罪（258 条）や私用文書毀棄罪（259条）の客体が「文書又は電磁的記録」と規定されていることと比較して、261 条に電磁的記録は含まれないからである。本判決は、「物」であるハードディスクが，本件の客体であるとしたが、本件で毀棄されたのは電磁的記録たるファイルであって，ハードディスク本体ではない。ハードディスク自体は，その本来の機能どおりに「読み出し」と「書き込み」を果たすのであり、改変されたのはファイルであると解すべきであろう。たしかに，本条の「損壊」については，物理的に物の全部または一部を害する場合のみならず，物の本来の効用を失わせる場合を含むとするのが本判例において引用されている判例でもあり、通説である（効用侵害説）[434]。これに対し，「損壊」とは，有形的な作用もしくは有形力の行使によって，物の全部または一部を物質的に破壊・毀損し，その結果としてその物の効用を害することをいうとする説も主張されている（物質的毀損説）[435]。物質的毀損説は，条文に忠実な解釈であり，本判決のような広い効用侵害説には「類推」の疑いがあるとされる[436]。事実、本判決における被告人は、「被告人の行為は、倫理的に問題のある行為ではあっても、ウイルス作成罪等の立法によって解決すべき問題であって、本件に器物損壊罪を適用することは刑法の類推解釈を認めるものであり、罪刑法定主義上許されない」と主張していたこと[437]も

---

433 浅田和茂「判例に見られる罪刑法定主義の危機」立命館法学 345・346 号（2012 年）13 頁、園田・前掲（注 425）108 頁。なお、森住・前掲（注 425）216 頁。

434 大塚仁・河上和雄・中山善房・古田佑紀『大コンメンタール刑法（第 3 版）第 13 巻』（青林書院、2018）807 頁（名取）。

435 たとえば、松宮孝明『刑法各論講義 〔第 5 版〕』（2020 年）324 頁以下。

436 浅田・前掲（注 433）14 頁。

437 同判決では「本年（2011 年一筆者注）7 月 14 日から施行された情報処理の高度化等に対処するための刑法等の一部を改正する法律によって新設された不正指令電磁的記録作成罪（刑法 168 条の 2）は、今

考慮すれば、想定事例の行為者に対する器物損壊罪の適用には疑問を残すことになるので、日本刑法においては、電子計算機損壊等業務妨害罪（刑法 234 条の 2）の適用可否を検討すべきである。

第 2 目　電子計算機損壊等業務妨害罪の適用可否

本罪の保護法益は、人の社会的活動としての業務遂行の円滑・安全であるとし、その構成要件は、人の業務に使用する電子計算機自体の損壊又はその電子計算機の用に供する電磁的記録の損壊という物理的な加害、人の業務に使用する電子計算機に虚偽の情報又は不正の指令を与えるという論理的な加害、あるいはこれらと同様の結果を惹起するその他の方法により、人の業務に使用する電子計算機に向けられた行為によって当該電子計算機の動作を阻害することである。また、本罪の行為客体は「人の業務に使用する電子計算機」としてよい[438]。

ここでいう「人の業務に使用する」とは、行為者以外の自然人、法人、法人格なき団体等であって、人が反復継続する意図のもとに行う経済的社会的活動たる業務の主体たる者がその業務に使用していることを意味している[439]。この点において、公務を「業務」に含むかについては見解が分かれるところであるが、すべての公務に使用される電子計算機は本条の客体になると解される[440]。

次に、「電子計算機」概念については定義規定が設けられていない以上、解釈を要するものとなる。しかし、単に「自動的に演算・データ処理を行う電子装置」と定義するのでは、例えばマイクロコンピュータを搭載する家電製品、自動販売機、電卓、電子辞書も本罪の客体と想定されることになる。しかし、本条が電子計算機に向けられた加害を手段とする新たな業務妨害行為をとらえて処罰しようとするのがその立法趣旨であるから、それ自体が自動的に情報処理を行う装置として一定の独立性をもって業務に用いられているもの，すなわちそれ自体が業務を左右するような判断，事務処理，制御等の機能を果たしている電子計算機といえるものに限定されるべきであり、およそ当該機器自体が自動的に情報処理を行う装置とはいえない家電製品や自動販売機は、一定の情報処理は行っているとはいえ，それ自体業務を左右するような判断、事務処理、制御等の機能を果たしていない電卓、電子辞書は本罪の客体としての「電子計算機」には当たらないとされる[441]。

後，本件のようなコンピュータウイルスを作成する行為にも適用されることになると推測され，法定刑も 3 年以下の懲役又は 50 万円以下の罰金と類似している。しかしながら、本件は、ウイルスによって被害者らのハードディスクを損壊したことを問題にしているのであって、ウイルス作成自体を処罰しようとするものではなく、両者は構成要件も保護法益も異なっている。したがって、不正指令電磁的記録作成罪の新設は、器物損壊罪の成否に影響しない」と判示して被告人の主張を退けている。

438 大塚ほか・前掲（注 434）247 頁（鶴田＝河村）。

439 大塚ほか・前掲（注 434）247 頁（鶴田＝河村）。

440 大塚ほか・前掲（注 434）248 頁（鶴田＝河村）。

441 大谷實『刑法講義各論〔新版第 5 版〕』（成文堂、2019 年）157 頁，山口厚『刑法各論〔第 2 版〕』（有

行為についてみると、第一類型である「損壊」とは、電子計算機や電磁的記録を物買的に変更，滅失させ，あるいは電磁的記録の消去などのようにその効用を害することとされ、第二類型である「虚偽の情報」とは，当該システムにおいて予定されている事務処理の目的に照らし，その内容が真実に反する情報のことであり，「不正な指令」とは，当該事務処理の場面において，与えられるべきでない指令のこととされる。また、第三類型である「その他の方法」は，電子計算機に向けられた加害手段であって，当該電子計算機の動作に直接影響を及ぼすような性質のものであることを要する。具体的には，電子計算機の電源を切断する，温度・湿度を急激に上下させるなどのような動作環境の破壊，通信回線の切断，入出力装置等の損壊，処理不能データの入力などが挙げられる[442]。

さらに本罪においては、「電子計算機に使用目的に沿うべき動作をさせず，又は使用目的に反する動作をさせ」るという結果，すなわち動作阻害という結果発生を必要としている。「使用目的に沿うべき動作」とは，電子計算機を設置して業務遂行のために使用する者が，具体的な業務遂行において当該電子計算機を使用して実現しようとしている目的に適合するような動作であって，例えば電子計算機がある条件下で一定の制御を行うという方法で機械制御に使用されている場合，そのような一定条件が与えられたときに行うこととされている制御の動作を意味する[443]。その場合の「動作」とは，電子計算機の機械としての働きをいい，具体的には，電子計算機の所定の機械制御を実行するため，必要とされている情報処理等のために行う出入力，演算等の働きのことである[444]。

最後に、「業務妨害」の要件が存在する。これは人が反復継続する意図で行う社会的活動である「業務」[445]を妨害することであるが、これには本罪は具体的危険犯であり、妨害の結果が現に生ずることまでは要せず、電子計算機に向けられた加害によって，実際にその「使用目的に沿うべき動作」をさせず，あるいは「使用目的に反する動作」をさせるという状態が発生し，これが業務を妨害するおそれのあるものでありさえすればよいとされる[446]。

本罪の構成要件上の問題として、AI が搭載された機器の内部データ変更やその損壊における事例について日本刑法に当てはめると、ドイツ刑法のそれとは異なり、業務に供される電子計算機でなくてはならないので、私用の介護ロボットなど、私的空間に属する AI 機器は本罪に該当しない。具体例をあげると、娯楽のために自動運転車（レベル 3 相当）を走行

斐閣、2012 年）166 頁、松宮・前掲（注 435）182 頁、浅田和茂『刑法各論〔第 2 版〕』（成文堂、2020 年）176 頁など。

442 大塚ほか・前掲（注 434）250 頁（鶴田＝河村）。

443 前田雅英編『条解刑法（第 4 版)』（弘文堂、2020 年）704 頁。

444 大塚ほか・前掲（注 434）250 頁（鶴田＝河村）。

445 その定義については、職業その他社会的活動に基づき継続して行う事務または事業をいうものと解される（大判大正 10 年 10 月 24 日刑録 27 輯 643 頁）。社会的活動としてなされる継続的な事務・事業は、広く本罪にいうところの「業務」に含まれ、商業・農業・工業等の経済的活動に限定されるわけではない。鎮目ほか・前掲（注 403）321 頁（鎮目）も参照。

446 大塚ほか・前掲（注 434）251 頁（鶴田＝河村）。

していたところ、ある者がハッキングにより同車と遠隔通信するサーバに侵入し同車の制御を奪った場合[447]や、介護の用に供する介護用ロボットにおける監視システムがハッキングされたことによりその制御を失い、被介護者の生命・身体に危険が迫る場合である。これらの事例において、本罪の成立は考えられず、これら AI 製品の制御を失わせたことに係る不正指令電磁的記録の作出行為のみが可罰的となるだろう。もっとも、医療現場で用いられるものなど、業務性が認められる AI 製品についてはこの限りではない。なお、業務を妨害するための手段がもっぱら不正指令電磁的記録によるものである場合、不正指令電磁的記録供用罪との罪数関係が問題となるが、本罪と不正指令電磁的記録供用罪の保護法益とは異なるため、事実関係が一個の行為と認められる場合には観念的競合になると考えられる[448]。

また、学習を行う AI 製品という観点での固有の問題は、前項のドイツ刑法での検討でも言及したように、第三者によってハッキングされた AI 製品が、その損壊ないしは虚偽の情報または不正な指令の供与等によりその使用目的に沿うべき操作をせず、またはその目的に反する動作をすることによってある者の業務を妨害した場合に、その損壊もしくは虚偽の情報または不正な指令の供与の原因がハッカーによるものなのか、AI の学習によるものなのかが不明であった場合、つまり妨害に至るまでの過程がブラックボックス化した場合に生じる。すなわち、ハッカーのハッキング以降になされた行為によって電子計算機等の損壊、虚偽の情報または不正な指令の供与等で人の業務妨害という結果を発生させたのかという因果関係の証明が問題であり、もしその原因が特定できないならばハッカーの行為と結果の因果関係が認められないことになるため、結果としてハッカーには電子計算機業務妨害未遂罪が成立するにとどまる。しかし、AI 製品の学習を隠れ蓑にしてハッカーが本罪の既遂の刑責をしようとする可能性も否定できない。そうだとすると、AI のブラックボックス性がサイバー攻撃に対する刑法上の評価が未遂罪にとどまってしまう結果となってしまうので、やはりそこで重要となるのが**説明可能な AI** の構想であり、想定事例におけるデータの損壊等が AI の学習によってもたらされたか否かを事後的に証明できるようにシステム構築・開発を行うことがサイバーセキュリティ上求められることになるだろう。ただし、私的空間に属する AI 製品に対するデータ変更・破壊行為は不正アクセス罪の適用のみが考慮されるにとどまり、不正指令電磁的記録を手段としてこの行為を行った場合に限り不正

---

447 この場合、自動車道における往来危険罪（道路運送法 100 条）の適用も考えられる。西貝吉晃「コネクティッドカーシステムに対するサイバー攻撃と犯罪」法律時報 91 巻 4 号（2019 年）49 頁以下参照。なお、この事例における「実行の着手」については同罪が具体的危険犯であることから、静止しているのではなく現に作動状態にある自動運転車にハッカーがその遠隔通信サーバに侵入し、同車の移動・停止を可能にした時点であるという（49 頁）。

448 大塚ほか・前掲（注 434）253 頁（鶴田＝河村）、杉山・吉田・前掲（注 385）90-91 頁など。なお、仮にこれらの場合で器物損壊罪も成立するならば、罪数関係としては、行為の単複により観念的競合か併合罪になると解すべきであるとされる（西田典之・山口厚・佐伯仁志編『注釈刑法 第 2 巻 各論（1）』（有斐閣、2020 年）554 頁（嶋矢）参照）。

指令電磁的記録供与罪（刑法 168 条の 2 第 1 項）が成立するにすぎない。

第 4 款 AI ソフトウェア・エージェントとコンピュータ詐欺

本項においては、AI を用いた資産運用を行うソフトウェア・エージェント（以下、「AI ソフトウェア・エージェント」とする）に対して不正なデータが用いられ、結果として AI ソフトウェア・エージェントの利用者に対して財産的損害が発生した事例を想定する。

第 1 項　AI ソフトウェア・エージェントとドイツ刑法におけるコンピュータ詐欺罪

コンピュータをネットワーク経由で攻撃する者は演算処理に影響を与えることで、コンピュータ詐欺を遂行しうる。データ処理設備の運営者の財産は、刑法 263 条 a の保護の対象である[449]が、この規定は、データ処理経過の違法な操作の場合、刑法 263 条についての欺罔の名宛人が欠けているためこの条文を適用することが困難であったという理由で可罰性の間隙を埋めるものとされ、1987 年に新たに規定されたものである[450]。

刑法 263 条 a の意味でのデータは刑法 202 条 a とは異なり、暗号化されたデータに限定される。そしてデータ処理とは、データを記録し、プログラムされたコードナンバーに従いそれらを結び付けることを通じて、一定の結果を得る電子データ処理システムにおける経過のことである[451]。ここで、学習能力のある AI 製品内の計算経過が刑法 263 条 a の意味でデータ処理と認められるか否かという問題が提起される。そもそもデータ処理経過は、入力データが処理された後に出力され、処理経過が「具体的な」結果に至ることにより特徴づけられる[452]。しかし現在では、従来の入力データはなく、学習能力を有する AI 製品がデータを収集し、あるいはどのデータを考慮するかを自ら決定することさえできるのではないかともいわれる[453]。その場合、入力データに応じて出力は変化するが、そのデータが直接システムに入力されるか、それとも AI の自立学習により自らデータを収集・評価してシステムに入力されたのかを事後的に区別することは困難である。そこで、刑法第 263 条 a の適用領域を、「複雑な、知性を補完する人工的な…知能」のみに限定されるべきであるという試みもあり[454]、この試みは、データ処理はプログラムデータ以外の新たな情報を吸収することが可能であり、この情報の区分化された分析と分類は、既存の保存されたデータや並行的に記録されているデータと比較またはリンクすることによって実行される必要があり、それゆえに直接に財産を減少させる機能も存在するという発想に基づく[455]。少なくともこのことから、学習能力を有する AI システムも刑法 263 条 a でのデータ概念に該当しうると結論づ

---

449 BT-Drs. 10/318, S. 12, 16 ff.; *Leupold/Glossner*, a.a.O. (fn.385), Teil 10. Rn. 140
450 *Barton,* Multimedia-Strafrecht Ein Handbuch für die Praxis, 1999, Rn. 21
451 *Leupold/Glossner*, a.a.O. (fn.385) Teil 10, Rn. 145.
452 MüKo-StGB/*Wohlers*, § 263a, Rn. 14.
453 *Günther*, a.a.O. (fn.176), S.234.
454 MüKo-StGB/*Wohlers,* § 263a. Rn. 15; その例として、ATM が引き合いに出される。Vgl. *Hilgendorf* Scheckkartenmißbrauch und Computerbetrug OLG Düsseldorf, NStZ-RR 1998, 137 JuS 1999.
455 MüKo-StGB/*Wohlers,* § 263a. Rn. 16

けられるだろう。

さらに、刑法第 263 条 a 第 1 項の 4 つの選択肢のいずれかを満たした上で、財産処分を惹起するようなデータ処理過程への影響が存在しなければならない[456]。この影響は、例えばプログラムの作出も含むプログラムの不正作成（ドイツ刑法第 263 条 a 第 1 項第 1 選択肢）に存在しうる[457]。ここでいう「不正」とは客観的な意味で理解され、そのプログラムが客観的にデータ処理タスクに適切に対処しているかどうかを考慮すべきであるため[458]、処分権限のあるユーザーの意思が介在しないことに注意しなければならない[459]。刑法 263 条 a 第 1 項第 2 第 2 選択肢は、不実または不完全なデータの使用（いわゆる入力操作）を処罰の対象としている[460]。データに含まれるその情報が現実と一致していない場合には不実性が、データから基礎に置く事情を十分に認識させない場合には不完全性が存在する[461]。データの無権限使用（刑法 263 条 a 第 1 項第 3 選択肢）では、正しいデータが権限者の意思に反して使用され、その使用がコンピュータ・システムではなく自然人に対して欺罔の性格を有することが前提となる[462]。第 4 選択肢の「経過へのその他の無権限干渉」は、AI 製品に関する構成要件的行為が、第 1 選択肢から第 3 選択肢と類似の結果不法と行為不法を有する場合[463]、例えば AI 製品のハードウェアの改造、プログラムのバグの悪用と関連する。ここでは受け皿構成要件として、特に新しい技術や未知の技術を把握することから、この類型が AI の分野でますます関連するものとなろう。

これらの犯罪遂行の 4 つの選択肢は、いずれも AI ソフトウェア・エージェントに対して大きな意味を持つ。例えば、電子商取引では、不正なデータを使用することや、オンラインショップでの価格表示の誤認で AI ソフトウェア・エージェントが「騙される」可能性があり、その結果 AI ソフトウェア・エージェント利用者の財産損害が発生しうる。また他人の署名を使用することも、データの無権限使用を充足する可能性がある。ここでは、第三者が当該 AI ソフトウェア・エージェントに偽の署名をして自らを認証するような場合が該当するだろう。第 4 選択肢は、他の類型と同置される、未知のあるいは新たな技術をカバーすることができるため、ロボット工学の範囲では関心が持たれうる。この類型では、学習能力を有する AI ソフトウェア・エージェントが、第三者から誤情報を与えられたられた場合に認められうる。そして、AI ソフトウェア・エージェントの利用者の財産が実際に減少（財産

456 *Leupold/Glossner* a.a.O. (fn.385)*,* Teil 10, Rn. 146.
457 *Marberth-Kubicki,* Computer- und Internetstrafrecht, 2009, Rn. *55.*
458 MüKo-StGB/*Wohlers*, § 263a. Rn. 22.
459 *Marberth-Kubick,* a.a.O. (fn.457) Rn. 56;
460 *Fische*r, a.a.O. (fn.181), § 263a, Rn. 7.
461 *Fischer*, a.a.O. (fn.181), § 263a. Rn. 7.
462 *Fischer*, a.a.O. (fn.181), § 263a, Rn. 10, 11. このことには争いがないので、技術的システムがまさに人間と比較可能でないということが議論される。「コンピュータ特有の」もしくは「詐欺特有の」解釈をめぐる争いについては、*Fischer,* a.a.O. (fn.181), § 263a. Rn. 9 ff.も参照。
463 *Fischer*, a.a.O. (fn.181), § 263a, Rn. 18.

損害）したことが本罪の成立には必要である[464]。このとき，資金移動の手続が不正に利用された場合などが定期的に存在するなど、資産に損害を与える（損害に相当する）リスクも財産的不利益を構成し本罪の対象となる[465]。

第 2 項　AI ソフトウェア・エージェントと日本刑法における電子計算機使用詐欺罪

キャッシュレス化がますます推進されている時代において、コンピュータによる処理のみが予定される取引形態もまた今後ますます拡大していくと予想されるところ[466]、もし行為客体が財物であれば、人の判断を介さずに財物の占有を取得する行為には窃盗罪（刑法 235 条）を適用しうるが、条文上客体が財物に限られる窃盗罪を、財産上の利益の場合に拡張することは許されないので、機械を不正に操作して人の判断を介在させずに財産上の利益を不正に取得する行為は、窃盗罪にも詐欺罪にも問えないこととなる[467]。本条は、このようなシステムを悪用する新たな財産侵害行為に対処するため，1987 年改正により設けられたものである。

本条の構成要件として、まず「財産権の得喪，変更に係る電磁的記録」とは，財産権の得喪，変更の事実又はその得喪，変更を生じさせるべき事実を記録した電磁的記録であって，一定の取引場面において，その作出（更新）により事実上当該財産権の得喪，変更が生じることとなるようなものをいう[468]。その財産権とは，金銭的価値を内容とする権利であって，債権，物権等がその典型であるとされ[469]、記録の作出等と事実上の財産権の得喪，変更との間の直接的あるいは必然的な関連性を要する[470]。次に、前段類型の「虚偽の情報」とは、電子計算機を使用する当該システムにおいて予定されている事務処理の目的に照らし，その

---

464 MüKo-StGB/*Hefendehl/Noll*, § 263a, Rn.179.

465 MüKo-StGB/*Hefendehl/Noll*, § 263a, Rn.179.

466 鎮目ほか・前掲（注 403）344 頁以下（鎮目）。

467 西田典之・山口厚・佐伯仁志編『注釈刑法 第 4 巻 各論（３）』（有斐閣、2020 年）317 頁以下（西田＝今井）によると、たとえば、ATM により他人の預金を自己の口座に振替送金する行為は、それだけではいまだ「財物」を取得したとはいえないために窃盗を構成しないし、自己の口座に他人の預金を不正に付け替えた後、いまだ現金化しない間に自動振替によって水道・ガスなどの利用料金が引き落とされた場合でも行為者は一度も現金を手にしていない」ので窃盗罪の成立は否定されるという。さらに、鎮目ほか・前掲（注 403）345 頁以下（鎮目）では、甲が A 銀行の係員に偽札を渡すことで 100 万円の入金処理をさせて預金債権を取得した場合にいわゆる 2 項詐欺罪（刑法 246 条 2 項） が成立するにもかかわらずインターネットバンキングを用いた場合には不可罰だとするのは不均衡であるという。インターネットバンキングを用いたとしても、その行為は、真実は経済的・資金的実体の伴う入金がないにもかかわらずそれがあったかのように装うものであり、これも「機械を欺いて」預金債権を取得する行為である。これは人に対する詐欺罪と同質の行為であるといえるから、詐欺罪と同等の処罰に値するものだという。

468 米澤・前掲（注 376）116 頁。

469 大塚ほか・前掲（注 434）13 巻 180 頁（和田）。

470 その理由については、鎮目ほか・前掲（注 403）346 頁（鎮目）参照。

内容が真実に反する情報」をいう（東京高判平成 5 年 6 月 29 日高刑集 46 巻 2 号 189 頁）。「不正な指令」とは当該事務処理の場面において，与えられるべきでない指令のことをいい、「不実の電磁的記録を作り」とは、真実に反する内容を，記録媒体上に電磁的記録を存在するにいたらしめることをいう。ここには記録をはじめから作り出す場合のほか，既存の記録を部分的に改変，抹消することによって新たな電磁的記録を存在するにいたらしめる場合も含まれる。後段類型の「財産権の得喪若しくは変更に係る虚偽の電磁的記録を人の事務処理の用に供する」とは、行為者が真実に反する財産権の得喪，変更に係る電磁的記録を他人の事務処理に使用される電子計算機において用い得る状態に置くことをいう。そして前段・後段に共通する要件としては不当利得があり、例えば、不実の電磁的記録を使用して銀行の預金元帳ファイルに一定の預金債権があるものと作為し，その預金の引出し，振替を行うことができる地位を得ることなど，事実上財産を自由に処分できる利益を得ること，不正に作出したプリペイドカードを利用して労務やサービスなど一定の役務の提供を受けること，料金の計算及び請求が行われることとなる課金ファイルの記録を改変して料金の請求を事実上免れることなどがこれに該当し、必ずしも実際に権利又は義務の得喪，変更が行われたことを要しない[471]。

第 4 項の冒頭で示した想定事例において、ドイツ刑法でのコンピュータ詐欺罪と比較すると不当利得要件を満たすか否かが問題となる。単なる財産損害にとどまる場合にはこの類型は該当せず、行為者に対し虚偽または不実のデータを供用したことについて、行為客体たる AI ソフトウェア・エージェントが業務に供するものである限りで電子計算機損壊等業務妨害罪の成立が考えられるが、私的用途で利用される AI ソフトウェア・エージェントの場合はその類型にも該当せず何らの犯罪も成立しないことになる。このことは、AI ソフトウェア・エージェントの利用者に対する財産的損害を与えることについて犯罪が成立しないことを意味するため、私見としては前項でも述べたように、電子計算機損壊等業務妨害罪の業務性要件を一般的利用にも拡張することで解決を図るべきであると思う。また、AI ソフトウェア・エージェントの学習に関する問題では、財産権の得喪もしくは変更に係る不実の電磁的記録の原因たる虚偽の情報もしくは不正な指令、ないしは財産権の得喪もしくは変更に係る虚偽の電磁的記録が人間の手によってもたらされたのか、それとも AI ソフトウェア・エージェントの学習によってもたらされたのかが不明な場合に問題となる。仮に行為者によってもたらされたデータが虚偽または不実のものであっても、不実の電磁的記録の作成が AI の学習によってなされたとすれば、行為者が AI の学習の結果、不実の電磁的記録を作成することを事前に認識していない限りで、本罪の行為類型を充足するとは言えず、行為者が不当利得を受けたとしても電子計算機使用詐欺既遂罪の成立は否定されてしまう。仮に行為者が不当利得を得る意思があったとしても、AI の学習過程があるか否かが既遂か未遂かの評価の分水嶺となってしまうことと、先述の電子計算機損壊等業務妨害罪やドイツ刑法におけるデータ変更・コンピュータ破壊罪のように、AI の「学習」によって行為者

[471] 大塚ほか・前掲（注 434）13 巻 187 頁（和田）。

の罪責が未遂減軽される可能性があることが問題だろう。ここでも重要なのが**説明可能な AI** の構想であり、行為者が当該 AI ソフトウェア・エージェントに与えた情報が不実の電磁的記録が作出したか否かを事後的に検証できるようなシステム構築を図ることが、AI の学習の不正な使用を防止すること、そして AI 学習に対する社会的信頼を確立するためにも必要なことである。

第 5 款 小括

本節では、主に学習機能を有する AI 製品がハッキングによるサイバー攻撃を受けた場合に想定される事例を、①AI 製品内に保存されているデータを取得した、②AI 製品の内部データを変更・破壊することによって利用者に一定の不利益が生じさせた、③AI ソフトウェア・エージェントに対し虚偽または不実の情報を供与してその利用者に財産的損害を与えた、という 3 つの類型に分類し、その手段行為のハッキング行為も含めてそれらの刑法上の評価を検討した結果、構成要件自体の解釈と AI の学習固有の問題があることが分かった。

まず、ハッキング行為についてドイツ刑法では、保護されるデータが日本における不正アクセス罪のようにアクセス制御機能を有しているか否かには関係なくドイツ刑法 202 条 a の対象となる。この点、日本刑法ではアクセス制御機能のないデータに対するアクセスは処罰の対象としていないことから、利用者・製造者の側でセキュリティを強化するようなシステム構築が必要であるといえる。データ取得について、ドイツ刑法では 202 条 b に所定される構成要件に該当する一方で、日本法では電気通信事業者法 179 条 1 項または有線電気通信法 14 条の罰則が関連することになる。データ変更・破壊にかかる利用者へ不利益を生じさせる行為については、ドイツ刑法では 303 条 b のデータ変更罪の構成要件に該当するが、この対象となるデータについてはその利用者にとって本質的に重要であるか否かが問題となる。その一方で、日本刑法では刑法 234 条の 2 所定の電子計算機損壊等業務妨害罪の成否が問題となるが、ここでは対象となる電子計算機がもっぱら業務に供するものでなければならないことに留意しなければならない。この点、私的空間で利用される AI 製品の場合、ドイツ刑法下では処罰の対象となりうるが、日本刑法のもとでは処罰の対象とならない。この比較から、AI 製品ひいては IoT 化された製品に対する保護の観点も踏まえて、業務性要件を保持しつつ、私的空間に属する電子計算機にもその範囲を拡げるべきではないかと考える。AI ソフトウェア・エージェントに対し虚偽または不実の情報を供与してその利用者に財産的損害を与えた行為について、ドイツ刑法では 263 条 a 所定のコンピュータ詐欺罪が、日本刑法ではで 246 条の 2 所定の電子計算機使用詐欺罪の成否が問題となる。前者では、当該行為についてドイツ刑法 263 条 a 第 1 項第 4 選択肢の受け皿構成要件に該当しうる一方で、後者では財産損害ではなく行為者の不当利得を求めることから、当該行為の構成要件には該当しないことになる。

そして、AI の学習のブラックボックス性は、上記のうちデータ変更・破壊とコンピュータ詐欺類型に関連する。前者においては、データの変更や削除という結果が、後者において

はその手段たる虚偽または不実の電磁的記録が、行為者（ハッカー）によるものなのか、それとも AI 自身の学習によるものかが不明確な場合、構成要件的結果が実現されたとしても因果関係が肯定されず未遂罪の適用にとどまり減軽の余地が残されてしまう。このアンバランスさの解消、そして因果関係の慎重な認定のため、**説明可能な AI** の構想に基づき、事後的にその因果関係を証明できるようなシステム構築を求めることが、AI 製品に対するセキュリティ上の保護、そして AI の学習に対する社会的信頼を確立することができるように思われる。

第 4 章　AI 製品開発に対する将来的な刑法上の規制

第 3 章でみてきたように、AI の利活用によっては現行刑法及び特別刑法の規制の射程が及ばなかったり、解釈上処罰（・制裁）範囲が拡張されたりする可能性があることが確認された。これらの議論はすでに人間の手によって作られ、これから発展する可能性のある AI に関する議論であることに留意しなければならない。この点において、そもそも人間社会にある一定の害を及ぼす可能性のある AI を創るべきでない、という見解もありうる。それは本稿で対象としてきた「弱い AI」の枠組を超える、「強い AI」にもその射程が及ぶという。それは、このような AI の開発に関し、事前抑制を意味するが、この点において刑法はどのような役割を果たすべきなのか。これについては、設計開発のみならず、研究自体も規制となりうる可能性を考慮しつつ、今後の AI 開発についての法的考察をまとめた Gaede の見解[472]を参照しながら検討を行う。

第 1 節　問題の所在―強い AI とその現状

強い AI の土台が、この瞬間にも世界中で真摯な試みとして自然科学的なものとして位置づけなければならない手法で研究されている[473]。また、人類の多くが AI の潜在能力を高める投資や能力に目を向けているように見えるという事実も十分に留意されないままに、経済、医療、軍事において、自らの将来を賭けることになるのではないかという不安の中で、実際に技術水準を変革するために多大な努力が尽くされている[474]。それと同時に、強い AI の研究により高性能のプラットフォームも作られつつあり、その一例として米国防総省の研究機関 DARPA は、すでに軍事的に利用可能なモデルやシステムの研究の一部でさえも AI が担うことになっている AI のプログラムを作成しており、それに対してエラーをする可能性があり、動作の遅い人間は「ゲーム・チェンジ AI」によって置き去りにされるという[475]。こうした見解に対する戦略的な対応は、強い AI は将来的にも不可能であると考えるこ

---

472 *Gaede*, Künstliche Intelligenz -Rechte und Strafen für Roboter? Plädoyer für eine Regulierung künstlicher Intelligenz jenseits ihrer reinen Anwendung, Nomos 2018.

473 *Gaede,* a.a.O. (fn.472), S.72 参照。ここには「フランケンシュタイン―保護領域の例外」がある。*HELG on AI*, *supra* (fn.95), p. 12 も見よ。

474 技術的革命については *Hilgendorf*, a.a.O. (fn.107) S. 99 f.も言及している。

475 それに対応する計画については https://www.darpa.mil/work-with-us/AI-next-campAIgn（最終アクセス

とにある。AI の研究者が尽力した、人間は計算可能で再構築する能力を持つ機械であるという争いのある想定は[476]、むしろ狂気の沙汰のようなものであるとともに、我々は、説明しえない霊感だけが知的生命体を創造することができると信じるために、自らや人類を買いかぶるマッドサイエンティストは明確なものなので不安を煽るものではないが、万一に備えて危険防止や処罰のための規制権限に触れる場合があるかもしれない[477]。

強い AI が実現するのはいつであるかという予測は問題にしないものの、もはや SF 映画の世界にそのテーマを委ねるべきではないと考えられる。とりわけ野心的な AI 研究を誇大妄想だと片付けようとする者でも、中心となる問題を誇大妄想家の制御をその手に委ねることになるが、実体的な法理解には、法が紙上の理想として公式化されるだけではないことが必要である。むしろ、我々は適切な配慮を講じることができるように、我々の法の形と執行可能性に対する実存的な危険に、たとえわずかな兆候であっても時宜に即して注意を向けるべきである[478]。

また、人間の自律性を高度に模倣したことによっても、重大な危険が生じうる。例えば、弱いながらも自己学習し、部分的には優れた能力を持っている AI が、誤解を伴った命令により、制御不能になったり、危険な状態になったりすることがある[479]。また、技術的な基盤が予測可能な形で改善されることで、自己意識を探求する研究者が、半ば輝かしい自己意識へのブレイクスルーを超克するために、意図的または無意識的に重大な効果を持つリスクを冒すという危険を高めるかもしれない。この関連で「Random Darknet Shopper」という AI も訴求力を持つ。その AI は、ダークウェブ上でドラッグと偽造パスポートをランダムに注文していたという[480]。システムは、確かに分別を持たないままではあるものの、予測不可能な危険な方法で「行動」することになることはこの点からも窺える。

人間と機械が共存する未来に向けて、危険防止や刑罰といった折り紙付きの手段への信頼は、決して将来への不渡手形をとなるべきではないが、そのためには個別利用の範囲外でも AI に目を向ける必要がある[481]。なぜなら、このようにして初めて包括的かつ技術的な危

2022 年 11 月 11 日）に記載される「米国国家安全保障のための新たなゲーム・チェンジ AI 技術」では,米国国防総省は 2030 年までに 1 億 6000 万ドルをさらに自律化した「未来戦闘部隊」に投じるとされる。

476 その立場を支持する石黒氏については *Eberl,*a.a.O.(fn.59) S. 321 も参照。

477 *Gaede,* a.a.O. (fn.472), S.72

478 *Gaede*, a.a.O. (fn.472), S.73

479 この例については *Bostrom,* Superintelligence - Paths, Dangers, Strategies, 2014, p.146. *Russell/Norvig,* Artificial Intelligence - A Modern Approach (3rd Edition), 2010, p.1010.後者は、例えば AI が首尾一貫しない、明確に定義されなかった道徳律から誤った結論を人間への対処について導出してしまう可能性を説明する。その一例として、その低い知能のために我々と同等の存在とはみなさない昆虫を殺しても良いという権限について挙げられている。

480 https://motherboard.vice.com/de/article/78kyz4/random-darknet-shopper-590　や https://motherboard.vice.com/de/article/kb7jma/kunstfreiheit-siegt-in-der-schweiz-duerfen-bots-drogen-im-darknet-kaufen-632.を参照（最終アクセス 2021 年 9 月 25 日）

481 例えば BT Drs.19/5880 を参照。

険分析が成功するからである。また、開発は世界中で行われ、容易には透明化されない共有財産が存在する状況を鑑みれば、すでにその中には第一の大きな挑戦が存在する。それは軍事的利用のために秘密となっていることや、企業秘密として自由に閲覧できない部分領域であるとされる[482]。

第 2 節　規制的措置

さらに、アナログもしくはデジタルでは操作できない手段を法執行に対して適用することも検討すべきであろう。人間の手によって操作されるという脅威は以前から存在していたが、いわゆるサイバーセキュリティの中ではこれについて争われている。とりわけ、AI の研究・構築の規制を検討することもまた必要であると思われる。

第 1 款 2010 年代における AI 製品開発・研究に対して考慮されてきた規制

弱い AI の利益を奪うような過剰な規制に陥りたくないならばどうすればよいか。これには「強い AI」を目指す、あるいはそれが当然だと思われる段階を含む研究が重要だと思われるが、こうした研究は産業用ロボットやサービスロボットのための、用途に方向づけられた規制や市場に方向づけられた規制、例えば、ロボットの利用であれば製品安全ガイドライン ISO 10377 や RL/2001/95/EG（欧州）、産業用ロボットであれば ISO 10218-1, ISO 10218-2:2011、サービスロボットであれば ISO 13849-2, ISO 18646-1:2016（ISO 18646-2/3/4,13482）といった規制を超えるものであることに留意すべきである。

では、このような AI 研究に必要な義務付けはどのようなものとなるのか。それは技術水準に従って、開発された AI がその評価に先立って研究空間や研究ネットワークを超えた影響を及ぼさないようにも義務付けをすることにあり、それには「封じ込め」の形式が必要とされる[483]。すなわち、新しい AI がテストされることなく他の技術システムに波及し、ないしはそれらとネットワーク化されないよう、物理的ないしはアナログ的に保証するべきという[484]。AI に対しては適用の動機から生じる、これまでの自発的な選好を超えた学習プロセスの導入を義務づけなければならないが、この場合、たとえ莫大な経済的投資をしたとしても、いまだ十分な安全性を確保できていない技術が排除されてしまうことは考えられないことであってもよい。もちろん、自律的な技術を多種多様に扱うため、いわゆる残存リスクからの絶対的な安全性を提供することはできないものの、極めてリスクのある研究に対して具体的で実現可能な注意基準を対置させることは可能である。すなわち、法の維持のためには、AI 研究に適切な法執行のインターフェースを、それに失敗した場合には、同等のメカニズムを再現することを追加で検討すべきである。このことについて欧州議会は、（AI

---

[482] *Gaede*, a.a.O. (fn.472), S.75.

[483] 自己学習する（産業用）ロボットにおける必要不可欠な「カプセル化」はますます困難となっていくであろうと認める *Stell/Krüger*, Lernen und Sicherheit in Interaktion mit Robotern aus Maschinensicht, in: *Hilgendorf/Günther*, Robotik und Gesetzgebung, Nomos, 2012, S.51, 61, 68 ff.も参照。

[484] *Gaede*, a.a.O.(fn.472) S.81.

を搭載する）ロボットは常に人間がいかなるときでも統制できるように構築されるべきだとさえ要請している[485]ことからも示される。

第 2 款 2020 年代に AI の製品開発に対して策定された国内外の規制

前項での議論がなされていた 2010 年代後半からさらに発展して、2020 年代には欧州連合、米国、中国で相次いで AI 開発に関する法的ガイドラインが策定された。以下、各国の AI 規制に関して概観し、それと日本の AI 規制を比較しながら、将来的な国際的レベルでの AI 開発の観点のもと望ましい規制となっているのか、そうでなければどのような問題点があるのかを抽出し、その解決策を提言したい。

第 1 項　欧州 AI 規制案（EU）

2021 年 4 月、欧州委員会は AI に関する規制法についての提案（以下、「欧州 AI 規制案」という。特に断りのない限り規制案の条文を引用する際には条文番号のみを記載する）である「Proposal for regulation of the European Parliament and of the Council: Laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts」を発表した[486]。この提案では「リスクベースのアプローチ」が強調されており、そのほかにもこの規制法案の特徴として、EU における既存の規制との調和、future-proof（将来の問題に対処できるような設計であること）、EU における統一的市場の確保、投資とイノベーションの促進といったことが強調されている。これらを見ると、コストとベネフィットの両方に配慮したビジネス的観点を重視したものとされる[487]。

この欧州 AI 規制案が我が国にもたらす影響としても留意しなければならないのは AI 法案の規制範囲が EU 域内だけでなく域外にも適用される点である[488]。欧州 AI 規制案は規制対象として次の 3 つを列挙する。すなわち、(a) 設立されたのが EU 域内であるか第三国であるかにかかわらず、EU において AI システムを市場に置き又はサービスを提供する提供者、(b) EU 域内に所在する AI システムの利用者 (c) AI システムが生み出すアウトプットが EU 域内で利用される場合における、第三国に所在する当該システムの提供者及び

---

485 Vgl. *Europäisches Parlament*, Resolution zu Zivilrechtlichen Regelungen im Bereich Robotik, P8_TA(2017)0051, Allgemeine Grundsätze bezüglich der Entwicklung der Robotik und der Künstlichen Intelligenz, 3. この要請は、主に意識を持たない AI に向けられたものと思われるが、欧州議会が除外していない自己意識型ロボットの場合、完全な制御可能性の要求は自由主義論の観点からは攻撃されうる。道徳的主体が承認される限り、原則的に自由が与えられなければならず、自由制限的、あるいは国家の連関の中で自由を定義するような規範の執行は種類や程度に応じて正当化されなければならない（Vgl. *Gaede,*a.a.O. (fn.476), S.79 fn.214）。

486 その原文（英文）は https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206（最終アクセス 2022 年 11 月 11 日）で閲覧可能である。

487 久木田水生「AI のリスクと倫理」第 36 回人工知能学会全国大会論文集（2022 年）1 頁。

488 北和樹「EU が目指す AI 社会のための規制法」立命館大学人文科学研究所紀要 131 号（2022 年）287 頁以下参照。

利用者である（2 条 1 項）。ここでの提供者とは、「有償か無償かにかかわらず、AI システムを開発するもしくは AI システムを市場に投入することまたは自身の名前もしくは商標で AI システムのサービスを提供することを目的として開発された AI システムを所有する、自然人、法人、公的機関、行政機関またはその他の機関」とし（3 条 2 号）、また利用者とは、AI システムが個人的な非専門的活動で使用される場合を除き、「自身の権限の下で AI システムを使用する自然人、法人、公的機関、行政機関またはその他の機関」とする（3 条 4 号）。

次に、規制案で制限を受ける AI のリスクについては 4 つの分類―「許容できないリスク」、「高リスク」、「限定リスク」、「最小限リスク」―が用いられている。

「許容できないリスク」に位置付けられ、禁止される AI の活動態様としては規制案 5 条 1 項に規定があり、本人または他の人の身体的または精神的危害を引き起こしたり、引き起こす可能性の高い仕方で個人の行動を実質的に歪めるために、個人の意識を超えたサブリミナル技術を用いる AI システムの市場投入、サービス提供、または使用（a 号）、本人または他の人の身体的または精神的危害を引き起こしたり、引き起こす可能性の高い仕方で、当該グループに属する個人の行動を実質的に歪めるために、年齢や身体的・精神的障害による特定グループの個人の脆弱性を悪用する AI システムの市場投入、サービス提供、または使用（b 号）、個人の社会的行動または既知もしくは予測された性格特性に基づき、(i)データが最初に生成または収集されたコンテキストとは関係のない社会的文脈における、特定の自然人またはそのグループ全体に対する不利益な、または望ましくない扱い、(ii)その社会的行動またはその重大さに対して不当なまたは不釣り合いな、特定の自然人またはそのグループ全体に対する不利益な、または望ましくない扱い、いずれかまたは両者をもたらすようなソーシャルスコアを用いて、特定の期間にわたる自然人の信頼性を評価または分類するための、公的機関による、または公的機関に代わっての AI システムの市場投入、サービス提供、または使用（c 号）、公的にアクセス可能な空間における法執行の目的での「リアルタイム」リモート生体識別システムの使用（d 号）の禁止が規定される[489]。この規制の実効性を担保するために、71 条 3 項で違反者に対し 3,000 万ユーロ以下の行政上の制裁金を、

[489] 小泉雄介「欧州 AI 規制案の概要」データ社会推進協議会データ倫理プライバシー研究 WG 資料（2021 年）https://www.i-ise.com/jp/information/report/2021/202106.pdf（最終アクセス 2022 年 11 月 11 日）8 頁ではそれぞれの類型の具体例としては以下のものをあげる。

(a) トラック運転手に可聴域でない音を聞かせて、健康かつ安全な範囲を超えて長時間運転させるようにする。AI はこのような効果を最大化する音域の発見に使用される。

(b) 音声アシスタントを組み込んだ人形が、楽しくクールなゲームを装って、未成年者に次第に危険な行動やチャレンジをするようにけしかける。

(c) AI システムが、医者予約の無断キャンセル、離婚など、親の取るに足らない、あるいは無関係な社会的な「不正行為」に基づいて、社会的ケアを必要としている子どもを特定する。

(d) ビデオカメラによってライブで撮影された全ての顔が、テロリストを特定するためにデータベースに対してリアルタイムでチェックされる。

違反者が企業である場合には、前会計年度の世界全体における売上総額の 6%以下の金額のうち、いずれか高い金額の行政上の制裁金を課す罰則が規定される。

また「高リスク」に位置付けられる AI とは、自然人の健康と安全、あるいは基本的な権利に高いリスクを生じさせるものとされている[490]。こういったシステムは特定の要件[491]を遵守し、事前の適合性評価（19 条）を受けているならば、欧州の市場に出すことが許される。高リスクの AI の利用について、提供者に対してはリスク管理システムの確立（9 条)、実装（10 条)、文書化（11 条)、動作中の記録保持（12 条)、利用者への透明性の確保（13 条)、健康、安全又は基本権に対するリスク防止又は最小化を目的とする人間による監視（14 条)、及びセキュリティの確保（15 条）を義務付ける（16 条以下)。また、付属書II—A に定められる製品の場合には、当該製品の製造者に対しても提供者と同様の義務を課す（24 条)。利用者に対しては、自ら入力データの管理を行う場合には、入力データが高リスク AI システムの意図された目的の点から見て関連性を有するものであることを確保する義務（29 条 3 項)、使用上の指示に基づいて高リスク AI システムの動作を監視する義務（29 条 4 項前段)、使用上の指示に従って使用した場合に AI システムがリスク[492]を示すことになる可能性があると考える理由がある場合には、提供者又は販売者に知らせるとともに、当該システムの使用を中止し、基本権を保護することを意図する EU 法上の義務の違反に該当する重大な事象又は機能不全を特定した場合にも、提供者又は販売者に知らせるとともに、AI システムの使用を中断する義務（29 条 4 項後段)、高リスク AI システムによって自動生成されたログが自らの管理下にある場合にそのログを維持する義務（29 条 5 項）を有する。これら要件または義務に違反した場合、2000 万ユーロ以下の制裁金、または違反者が企業の場合は、直前の会計年度における世界全体における売上総額の 4%以下の金額、もしくはいずれか高額の方の制裁金の賦課という罰則を定める（71 条 4 項）[493]。

「限定リスク」もしくは「最小限リスク」に位置付けられる AI としては、人と交流することを目的とした AI システム、感情を認識するために使用されたり、生体認証データに基づいて（社会的な）カテゴリーとの関係性を判断したりするシステム、存在する人・モノ・

---

490 その内容は規制案 6 条 1 項と 2 項および付属書Ⅱ、付属書Ⅲに記載がある。付属書Ⅱ—A では、機械、玩具、娯楽用船舶、昇降機、医療機器などに関する EU 規則内で対象とする製品に付属書Ⅱ—B では、航空機、鉄道、自動車などに関する EU 規則内で対象とする製品に第三者適合性評価（43 条）の義務を提供者に課し（19 条)、付属書Ⅲでは自然人の生体識別・分類、重要なインフラの管理・運営、教育・職業訓練、雇用・労働者管理、及び自営業へのアクセス、重要な民間・公共のサービス及び給付へのアクセス及び享受、法執行など高リスクに該当する類型を列挙している。

491 規制案 8 条から 15 条にかけてその内容が具体化される。

492 そのリスクの内容は Article 3, point 19 of Regulation (EU) 2019/1020 を引用する。そこでは、合理的かつ許容可能と考えられる程度を超えて、一般人の健康および安全、職場における健康および安全、消費者の保護、環境、公安および該当する法令で保護されるその他の公益に悪影響を及ぼすものと定義される。

493 ただし、データによるモデル学習を伴った技法を利用する「高リスク AI」の開発要件である 10 条に違反した場合は 71 条 3 項の制裁の対象となる。

場所・その他の存在に酷似し、本物または真実であると誤解させる（「ディープフェイク」）画像、音声、またはビデオコンテンツを生成または操作する AI システムを指す[494]。これら AI システムに関する法的義務として、AI システムと相互作用をしている人々に、それが明白でない限り、チャットボットなどその旨を通知すること、感情認識システムや生体カテゴライゼーションシステムの対象となる人々にその旨を通知すること、表現の自由などの基本的権利の行使や、公共の利益の理由からディープフェイクが必要な場合を除いてディープフェイクコンテンツにラベルを付けることが規定される（52 条）。また、AI によってサポートされるゲームアプリケーションやスパムフィルタ機能 AI などは「最小限リスク」として位置づけられ、この種類の AI の任意の適用を促すことを目的とした行動規範について、当該 AI システムの意図された目的を考慮して、当該要件の遵守を確保する適切な手段である技術上の仕様及びソリューションを基礎として作成することを促進するにとどめる（69 条）。ただし、これら「限定リスク」や「最小限リスク」に該当する AI であっても、71 条 4 項の文言に従えば「高リスク AI」の場合と同様の制裁が課されることになることに留意しなければならない。

これら 4 類型のうち、「許容できない AI リスク」を除くもののリスクを示す AI システムに関しては、市場監視機関が、当該評価の過程で、AI システムが本規則に定める要件及び義務を遵守していないことを発見した場合、その市場監視機関は、AI システムをして当該要件及び義務を遵守させるために、AI システムを市場から取り下げるために、又は AI システムをリコールするために、リスクの性質に比例した当該市場監視機関が定め得る合理的な期間内に、適切な全ての是正措置を講じるよう、遅滞なく関係する事業者に要求するものとし（65 条 2 項）、事業者が適切な是正措置を講じない場合には市場監視機関が製品を当該市場から取り下げるための、又は製品をリコールするための適切な暫定措置を講じるとする（65 条 5 項）という制裁規範がさらに存在することも見過ごしてはならない。

この規制に関して EU の産業界からは，企業の負担が増加することに関する懸念が規制案発効後すぐに表明され[495]、イノベーションの阻害となることが示されている[496]。また、欧州の機械電気電子金属加工産業連盟（Orgalim）は、AI システムという定義が不明確なのでその定義をより明確化するとともに、産業用 AI は高リスクとみなされないことを保証することを求めること、適合性評価の義務化は企業の負担を増やし，安全性を高めることには必ずしもつながらないのではないのかという懸念を示し[497]、経団連の声明においても、「罰則の対象が広範に及び、また罰金額が非常に高額であることは、欧州市場における企業の活動

494 北・前掲（注 488）294 頁。

495 その具体的な内容については、寺田麻佑・板倉陽一郎「欧州（EU）における 2021 年 AI 規制法案をめぐる各種意見と EU の対応の検討」情報処理学会研究報告 22 号（2022 年）2 頁以下も参照。

496 See, BCS: New EU AI regulations demand a 'fully professionalised tech industry' - institute for IT. 2021, Apr 22.

497 *Orgalim*, European Regulation on Artificial Intelligence – Orgalim calls for legal clarity and workability, 21 April, 2021 (https://orgalim.eu/news/european-regulation-artificial-intelligence-orgalim-calls-legal-clarity-and-workability 最終アクセス 2022 年 11 月 12 日)

を過度に委縮させる恐れがある。そこで、違反の類型や内容、得られた便益の大きさ、違反の悪意の有無などに応じて、適切なペナルティを定めるべき」という。私見としても、当該規制の文言解釈からすれば、そもそも 71 条の表題が罰則（penalties）であること[498]と、制裁金（administrative fine）が課される要件が 5 条所定の「許容できない AI リスク」の利用に該当するのみならず、「高リスク AI」の利用における他の義務やその他「限定リスク」や「最小限リスク」AI に課せられうる義務まで制裁金の対象となる。さらに、市場監視機関がそれらの義務の不遵守を発見した際にはその事業者に対して製品のリコールのみならず市場からの取下げを命じることもできることに鑑みれば、これら規制は営業の自由を侵しかねない強い制裁規範であり刑罰的性格を帯びたものといってもよいだろう。しかもこのような「財産刑」の対象となる AI のリスク評価がもっぱら市場監視機関に委ねられており、さらには抽象的なリスク段階での規制であることも考慮すればこのような規制の運用自体をより慎重にしなければならない。

第 2 項　AI 権利章典（米国）

2022 年 10 月 4 日、米国ホワイトハウス科学技術政策局（OSTP: Office of Science and Technology Policy）は AI の開発に考慮すべき原則をまとめた「AI 権利章典のための青写真」（Blueprint for an AI Bill of Rights: A Vision for Protecting Our Civil Rights in the Algorithmic Age）を公表した。ここでは、患者の治療に役立つはずのシステムが、安全でないこと、効果がないこと、あるいは偏ったものであること、雇用や信用に関する判断に使われるアルゴリズムが不公平を反映・再現したり、新たな有害な偏見や差別を埋め込んだりしていること、ソーシャルメディアにおける無制限のデータ収集がプライバシーを損ない、本人の認識や同意なしに人々の活動を広く追跡するために利用されていることを問題の根底に置き、そのような公民権に対する脅威からすべての米国民を守り、その最高の価値を強化する方法でテクノロジーを活用する社会のための指針を、①安全かつ効果的なシステム(Safe and Effective System)、②アルゴリズム的差別からの保護(Algorithmic Discrimination Protections)、③データ・プライバシー(Data Privacy)、④通知と説明(Notification and Explanation)、⑤人間への代替、考慮、予備的措置(Human Alternatives, Consideration, and Fallback)の 5 つの原則にまとめたものである[499]。

①「安全かつ効果的システム」では、システムは多様なコミュニティやステークホルダ、専門家と協議の上、開発を行うものとし、システムを配備する前に試験を行い、リスクを特定・軽減し、システムの監視を行う。これらの保護措置の結果として、場合によってはシス

498 なお、本規制案のドイツ語版（https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=CELEX:52021PC0206&from=EN（最終アクセス 2022 年 11 月 13 日））での 71 条の表題は「Sanktionen」となっている。

499 OSTP, Blueprint for an AI Bill of Rights: A Vision for Protecting Our Civil Rights in the Algorithmic Age, October 14th, 2022.

テムの配備中止や削除もあり得るとする[500]。その実践のために、AI が(a) 合法的かつ国家の価値を尊重し、(b) 目的を持ちパフォーマンスを重視し、(c) 正確かつ信頼可能で効果的に、(d) 安全、堅牢で弾力性があり、(e) 理解可能で、(f) 責任があり追跡可能な、(g) 定期的に監視され、(h) 透明性を有し、(i) 説明責任を有するものであることを求める。

②のアルゴリズム的差別からの保護は、システムが人種、肌の色、民族、性別、宗教、年齢、国籍、障害、退役軍人の地位、遺伝情報、または法律で保護されているその他の分類に基づいて人々を不当に異なる扱いや影響を与え、このようなアルゴリズムによる差別は法的保護に違反する可能性を示唆しつつ、自動化システムの設計者、開発者、配備者は、アルゴリズムによる差別から個人やコミュニティを保護し，公平な方法でシステムを使用・設計するために，積極的かつ継続的な措置を講じるものとする。この保護には，システム設計の一環としての積極的な公平性評価，代表的なデータの使用と人口統計的特徴に対する保護，設計と開発における障害者のアクセシビリティの確保，配備前および継続中の格差テストと緩和，明確な組織の**監視**が含まれる必要がある[501]。

③のデータ・プライバシーでは、個人の合理的な期待に沿うもので、厳密に必要なデータのみを収集したうえで、システムの設計者、開発者、配備者は個人からの許可を取得し、データの収集、使用、アクセス、移転、削除に関する個人の決定を尊重する。個人の同意を求める際は、簡潔で、平易な言葉で理解できる内容にし、健康や仕事などに関わる機微なデータについては、**継続的な監視とモニタリング**をつうじてより強い保護措置を講じるものとする[502]。

④「通知と説明」では、システムの設計者、開発者、配備者は、システム全体の機能と自動化が果たす役割、そのようなシステムが使用されていることの通知、システムに責任を持つ個人・組織、明確で適時かつアクセス可能な結果の計画を明確に説明する文書を広く一般に提供する。これらの情報は最新の状態に保ち、重要な使用例や主要機能の変更についてはシステムの影響を受ける人々に通知するものとする。自動化システムは、技術的に有効で、利用者及びシステムを理解する必要のあるオペレータ等にとって有意義かつ有用であり、かつ文脈に基づくリスクレベルに適合した説明を提供しなければならず、これらシステムに関する要約情報を平易な言葉で記載した報告書、および通知と説明の明確性評価と質的評価を、可能な限り公表することも求める[503]。

⑤「人間による代替、考慮、予備的措置」では、システムから影響を受ける個人が必要に応じてオプトアウトし、人間による代替手段を選ぶことができるようにする。その適切性については、与えられた文脈における合理的予期に基づき、幅広いアクセス性を確保し、特に有害な影響から公衆を保護することに重点を置いて決定されなければならない。さらに、システムの失敗やエラーが起きた場合などに人間による考慮と予備的措置による救済を受け

500 OSTP, Blueprint, *supra* (fn.499), p.15.
501 OSTP, Blueprint, *supra* (fn.499), p.23.
502 OSTP, Blueprint, *supra* (fn.499), p.30.
503 OSTP, Blueprint, *supra* (fn.499), p.40.

られるようにする。ここではアクセス可能で、公平で、効果的で、維持され、適切なオペレータの訓練を伴うべきであり、一般大衆に無理な負担を強いないようにすることが求められる[504]。

これら 5 つの原則は、自動化システムの構築、展開、ガバナンスにおいて、市民の権利を保護し、民主的な価値を促進する政策と実践の開発を支援することを目的とするものにとどまり、それ自体は拘束力を持たず、既存の法令、規制、政策、国際文書に取って代わるものでも、それを修正するものでも、その解釈を指示するものなく、一般市民や連邦政府機関に対して上記原則の遵守を義務付けるものではないとしている[505]。むしろ、Executive Order 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government (December 2020)という大統領令など、既存の政策や規制に従うことが原則であることを留保している。

まとめると、この原則自体には法的拘束力はなく、制裁規範も有していないため、これら諸原則を製造者等に課せられた（刑）法的な注意義務として援用するのは少々困難なようにも思われる。しかし、この「AI 権利章典」が「青写真」の段階から実際に法的拘束力を有する「AI 権利章典」へと昇華した場合には、ここで掲げられている諸原則が AI 製品開発に関与する製造者に課せられる法的義務となり、注意義務の認定に資するものとなるだろう。

第 3 項　「新時代の人工知能倫理規範」（中国）

2021 年 9 月 25 日に中国の「新世代人工知能のガバナンスに関する国家専門委員会」は、人工知能のライフサイクル全体に倫理を統合し、AI 関連の活動に従事する自然人、法人、その他の関連機関等に倫理的なガイドラインを提供することを目的とした「新時代の人工知能倫理規範（新一代人工智能伦理规范）」（以下、「倫理規範」という）を公表した[506]。この倫理規範は、プライバシー、偏見、差別、公正性など、現在のコミュニティの倫理的懸念を十分に考慮し、テーマ別の調査、重点的な起草、協議を経て、一般規定、特定の活動に関する倫理規範、組織的実施事項に分類して作成された。倫理規範では、「人間の福祉の増進」「公正と正義の推進」「プライバシーとセキュリティの保護」「制御性と信頼性の確保」「責任の強化」「倫理意識の向上」という 6 つの基本的な倫理要件を定めると同時に、AI の管理、研究開発、供給、利用など特定の活動に対する 18 の具体的な倫理的要求事項を提案する。

このうち、AI 製品に関与する主体に関連する義務としては以下のようなものが挙げられる。研究開発者に関しては、技術研究開発のあらゆる側面に AI 倫理を統合することを率先して行い、意識的に自己検閲を行い、自己管理を強化し、倫理・道徳に反する AI 研究開発を自制する意識の強化（10 条）、データの収集、保存、使用、処理、伝送、提供及び開示の

504 OSTP, Blueprint, *supra* (fn.499), p.46.
505 OSTP, Blueprint, *supra* (fn.499), p.2.
506 中华人民共和国科学技术部・前掲（注 297）。この倫理規範については https://www.most.gov.cn/kjbgz/202109/t20210926_177063.html で閲覧可能である（最終アクセス 2022 年 11 月 28 日）

過程において、データ関連の法律、基準及び規範を厳守し、データの完全性、適時性、一貫性、標準化、正確性等データの品質向上（11 条)、アルゴリズムの設計・実装・応用において、透明性、解釈性、理解性、信頼性、制御性を高め、AI システムの回復力、自己適応性、反干渉性を強化し、検証性、監査性、監督性、追跡性、予測性、信頼性の実現（12 条)、データ収集やアルゴリズム開発において、倫理的審査を強化し、差別的主張を十分に考慮し、データやアルゴリズムの偏りの可能性を回避し、AI システムの普遍性、公平性、無差別性の実現（14 条）の義務がある。

次に製造（販売）者については、市場参入、競争、取引などの活動に関する各種規則を厳格に遵守し、市場秩序を積極的に維持し、AI の発展に資する市場環境を整備し、データ独占、プラットフォーム独占などで秩序ある市場競争を損なうことを控え、いかなる方法によっても他の主体の知的財産権を侵害することを禁止することを目的とする市場ルールの尊重（14 条)、AI 製品・サービスの品質監視と利用評価を強化し、設計や製品の欠陥などによる個人の安全、財産の安全、利用者のプライバシーの侵害を回避し、品質基準を満たさない製品・サービスの運用、販売、提供は行わないこと目的とする品質管理強化（15 条)、製品及びサービスにおける AI 技術の使用について、利用者に明確に伝え、その機能と制限を明らかにし、利用者の情報及び同意の権利を保護することを目的とする利用者の権利・利益保護（16 条)、緊急時のメカニズムや損失補償のスキームや手段を研究・開発し、AI システムを適時に**監視**し、利用者からのフィードバックに適時に対応・処理し、システム障害を適時に防止し、法律や規則に従って AI システムに介入する関連主体を支援し、損失を減らしリスクを回避する準備を整えることを目的とする緊急時の保護強化（17 条）を規定する。

さらに使用規定として、製造者に対しては AI 製品とサービスの使用前のデモンストレーションと評価を強化し、AI 製品とサービスがもたらす利益を十分に理解し、すべてのステークホルダの合法的権益を十分に考慮し、経済繁栄、社会進歩、持続可能な発展を促進するという善意の利用の促進（18 条）や AI の倫理的ガバナンスの実践に積極的に参加し、関連するテーマに適時にフィードバックし、AI 製品やサービスを利用する過程で見つかった技術的な安全性の陥穽、政策や規制の空白、規制の遅れなどの問題解決を支援すること（21 条）を、製造者のみならず利用者を含みうるものとして、AI 製品・サービスの適用範囲と悪影響を十分に理解し、関連する対象者の AI 製品・サービスを使用しない権利を効果的に尊重し、AI 製品・サービスの不適切な使用や濫用を避け、意図せずに第三者の正当な権利や利益を損なわないようにするという AI の誤用・濫用の回避（19 条)、法令・倫理・基準・規範に適合しない AI 製品・サービスの利用の禁止、AI 製品・サービスの利用による違法行為の禁止、国家の安全、公共の安全、生産の安全を脅かすことの禁止、公益の毀損等の禁止という AI の違法な利用の禁止（20 条)、AI 製品とサービスを安全に使用し、効率的に活用するために、AI に関連する知識を積極的に学び、運用・保守・緊急時の処理などに必要なスキルを率先して習得する（22 条）ことを定める。

この「倫理規範」は他の AI に関する規制とは異なり、製造開発に限定せず、利用者も含

め AI 製品に関与しうる主体の義務を明文化したものである。特に、AI 製品に起因する事故の刑事過失責任を検討する際に注意義務の確定の手掛かりとなるものとして、製造者ならば 15 条が設計上の義務・製造上の義務と、16 条が指示上の義務、17 条・21 条が製品監視義務と調和する。そして、利用者の注意義務として 19 条・20 条の内容が関連する。ただし、その具体的内容については条文上明確でないところは多いものの、これら法的義務をつうじて注意義務違反を認定することが望ましいだろう。もちろん、注意義務違反自体が刑事責任を生ぜしめるわけではなく、注意義務違反と結果の因果関係も求められる。その認定の障壁となりうる AI のブラックボックス性を可能な限り透明化する（＝説明可能な AI）ことも開発者に対して 12 条のように法的に義務付けることで、因果関係の適切な立証に資するものとなり、AI 製品にかかる事故事例における刑事責任の負責が過度なものにもならず、間隙となる状況も防ぐことができるように思われるため、これは我が国においても非常に示唆に富む AI 規範となるだろう[507]。

第 4 項　「AI 開発ガイドライン」・「AI 利活用ガイドライン」（日本）

我が国における AI 開発・利活用に関する法的原則としては、総務省による「国際的な議論のための AI 開発ガイドライン案」[508]（2017 年、以下「AI 開発ガイドライン」という）及

---

[507] この倫理規範から AI 製品の事故事例における刑事責任を論じた中国の文献としては、曾粤兴・高正旭「论人工智能技术的刑法归责路径」治理研究(2022 年第 3 期)113 頁以下がある。同文献では、ネットワークへの依存度が高い AI には、ネットワークセキュリティに関する包括的な法規範が必要であるとし、AI を支えるアルゴリズムの安全性を先行法規のみならず刑法のレベルでも担保すべきという。そこで、AI（アルゴリズム）の安全性を保護法益とし、魏东「人工智能算法安全犯罪观及其规范刑法学展开」政法论丛(2020 年第 3 期)を引用しながら「安全基準を満たさない AI 製品を設計・製造・販売・使用する罪、AI 武器を違法に設計・製造・所持・取引・運搬・使用する罪、AI 製品のアルゴリズムや使用方法を無断で改変する罪、AI 濫用罪、AI 騒乱罪」という 5 種類のアルゴリズムの安全に危害を加える犯罪の新設を提言するという試みがあり、さらに 17 条を参照しつつ、AI を供給する主体は、「緊急メカニズムや損害賠償計画・対策を検討・策定し、AI システムを適時に監視し、ユーザーからのフィードバックに適時に対応・処理し、システム障害を適時に防止し、関連主体が法律に従って AI システムに介入し、損失を軽減しリスクを回避できるように支援する準備をする」義務を負い、この義務に反してアルゴリズムの安全性を著しく侵害する行為を規制するために、刑法に新たな犯罪を創設する必要があるという（123 頁）。

[508] AI ネットワーク化の健全な進展及び AI システムの便益の増進に関する原則として、①連携の原則：開発者は、AI システムの相互接続性と相互運用性に留意すること、主に AI システムのリスクの抑制に関する原則として、② 透明性の原則：開発者は、AI システムの入出力の検証可能性及び判断結果の説明可能性に留意すること、③ 制御可能性の原則：開発者は、AI システムの制御可能性に留意すること、④安全の原則：開発者は、AI システムがアクチュエータ等を通じて利用者及び第三者の生命・身体・財産に危害を及ぼすことがないよう配慮すること、⑤セキュリティの原則：開発者は、AI システムのセキュリティに留意すること、⑥プライバシーの原則：開発者は、AI システムにより利用者及び第三者のプライバシーが侵害されないよう配慮すること、⑦倫理の原則：開発者は、AI システムの開発において、人間の尊厳と個人の自律を尊重すること、そして主に利用者等の受容性の向上に関する原則として、⑧利用者支援の原則：開発者は、AI システムが利用者を支援し、利用者に選択の機会を適切に提供することが

び「AI 利活用ガイドライン～AI 利活用のためのプラクティカルリファレンス～」[509]（2019 年、以下「AI 利活用ガイドライン」という）が存在する。

このうち、本論文で想定してきた事例で関連するのは、（１）AI 製品の監視・管理の観点においては「AI 開発ガイドライン」の⑤制御可能性の原則と「AI 利活用ガイドライン」の①適正利用の原則であり、（２）AI のブラックボックス性への対応として「AI 開発ガイドライン」の②透明性の原則と「AI 利活用ガイドライン」の⑨透明性の原則、そして（３）人間への責任帰属という観点からは「AI 開発ガイドライン」の⑨アカウンタビリティの原則、「AI 利活用ガイドライン」の⑩アカウンタビリティの原則である。以下、その具体的内容を確認する。

AI 製品の監視・管理に関して、「AI 開発ガイドライン」の③制御可能性の原則は、開発者に対して「AI システムの制御可能性に関するリスクを評価するため、あらかじめ検証及び妥当性の確認を行うよう努めることが望ましい」とし、「こうしたリスク評価の手法としては、社会において実用化される前の段階において、実験室内やセキュリティが確保されたサンドボックスなどの閉鎖空間において実験を行うこと」、そして「制御可能性を確保するため、採用する技術の特性に照らして可能な範囲において、人間や信頼できる他の AI による監督（監視、警告など）や対処（AI システムの停止、ネットワークからの切断、修理など）の実効性に留意することが望ましい」という[510]。一方で「AI 利活用ガイドライン」の①適正利用の原則では、「AI サービスプロバイダ及びビジネス利用者は、開発者等からの情報提供や説明を踏まえ、AI を利活用する際の社会的文脈に応じ、AI を利用する目的、用途と AI

可能となるよう配慮すること、⑨アカウンタビリティの原則：開発者は、利用者を含むステークホルダに対しアカウンタビリティを果たすよう努めること、という 9 つの原則を定める。

[509] ①適正利用の原則：利用者は、人間と AI システムとの間及び利用者間における適切な役割分担のもと、適正な範囲及び方法で AI システム又は AI サービスを利用するよう努める。②適正学習の原則：利用者及びデータ提供者は、AI システムの学習等に用いるデータの質に留意する。③連携の原則：AI サービスプロバイダ、ビジネス利用者及びデータ提供者は、AI システム又は AI サービス相互間の連携に留意する。また、利用者は、AI システムがネットワーク化することによってリスクが惹起・増幅される可能性があることに留意する。④安全の原則：利用者は、AI システム又は AI サービスの利活用により、アクチュエータ等を通じて、利用者及び第三者の生命・身体・財産に危害を及ぼすことがないよう配慮する。⑤セキュリティの原則：利用者及びデータ提供者は、AI システム又は AI サービスのセキュリティに留意する。 ⑥プライバシーの原則：利用者及びデータ提供者は、AI システム又は AI サービスの利活用において、他者又は自己のプライバシーが侵害されないよう配慮する。⑦尊厳・自律の原則：利用者は、AI システム又は AI サービスの利活用において、人間の尊厳と個人の自律を尊重する。⑧公平性の原則：AI サービスプロバイダ、ビジネス利用者及びデータ提供者は、AI システム又は AI サービスの判断にバイアスが含まれる可能性があることに留意し、また、AI システム又は AI サービスの判断によって個人及び集団が不当に差別されないよう配慮する。⑨透明性の原則：AI サービスプロバイダ及びビジネス利用者は、AI システム又は AI サービスの入出力等の検証可能性及び判断結果の説明可能性に留意する。⑩アカウンタビリティの原則：利用者は、ステークホルダに対しアカウンタビリティを果たすよう努める、と 10 の原則を定める。

[510] 総務省「国際的な議論のための AI 開発ガイドライン案」（2017 年）8 頁以下。

の性質、能力等を適切に認識した上で、AI を適正な範囲・方法で利用すること」、さらに「AI サービスプロバイダは、AI サービスの公平な条件による利用を確保するとともに、必要な情報を適時に提供することが期待される」[511]という。

AI のブラックボックス性への対応として、「AI 開発ガイドライン」の②透明性の原則では、「本原則の対象となる AI システムとしては、利用者及び第三者の生命、身体、自由、プライバシー、財産などに影響を及ぼす可能性のある AI システムが想定される」ことを前提に、「開発者は、AI システムに対する利用者を含む社会の理解と信頼が得られるよう、採用する技術の特性や用途に照らし合理的な範囲で、AI システムの入出力の検証可能性及び判断結果の説明可能性に留意することが望ましい」[512]とし、「AI 利活用ガイドライン」の⑨透明性の原則では、「AI サービスプロバイダ及びビジネス利用者は、AI の入出力等の検証可能性を確保するため、入出力等のログを記録・保存すること」、「ログの記録・保存に当たっては、利用する技術の特性及び用途に照らして、ログの記録・保存の目的、ログの取得・記録の頻度等について考慮することが期待される」という[513]。

そして、AI 製品にかかる責任の観点で、「AI 開発ガイドライン」の⑨アカウンタビリティの原則では、「開発者は、AI システムへの利用者や社会の信頼を得られるよう、自らの開発する AI システムについてアカウンタビリティを果たすことが期待される。具体的には、利用者に AI システムの選択及び利活用に資する情報を提供するとともに、利用者を含む社会による AI システムの受容性を向上するため、開発者は…利用者等に対し自らの開発する AI システムの技術的特性について情報提供と説明を行うほか、多様なステークホルダとの対話を通じて様々な意見を聴取するなど、ステークホルダの積極的な関与（フィードバック）を得るよう努めること」、そして「自らの開発する AI システムによってサービスを提供するプロバイダ等と情報を共有し、協力するよう努めることが望ましい」[514]といい、「AI 利活用ガイドライン」の⑩アカウンタビリティの原則では「AI サービスプロバイダ及びビジネス利用者は、人々と社会から AI への信頼を獲得することができるよう…消費者的利用者、AI の利活用により影響を受ける第三者等に対し、利用する AI の性質及び目的等に照らして、それぞれが有する知識や能力の多寡に応じ、AI システムの特性について情報提供と説明を行うことや、多様なステークホルダとの対話を行うこと等により、相応のアカウンタビリティを果たすよう努めることが期待される」[515]とする。

これらガイドラインの策定以来、国際的に新たな AI 開発や利活用に関するガイドライン・綱領が策定されている状況に鑑みて、我が国でも新たなガイドラインの策定の議論がなさ

---

511 総務省「AI 利活用ガイドライン～AI 利活用のためのプラクティカルリファレンス～」（2019 年）13 頁。

512 総務省・前掲（注 510）8 頁

513 総務省・前掲（注 511）24 頁。

514 総務省・前掲（注 510）11 頁以下。

515 総務省・前掲（注 511）25 頁。

れている[516]。そこでは 22 の尊重すべき価値[517]を確認しており、2019 年当時のガイドラインの国際比較と比べて、堅牢性、責任[518]、追跡可能性[519]、モニタリング・監査、ガバナンス、その他（コスト・効果測定）の 6 つの新たな項目の追加が検討された。

これらガイドラインの規定で、透明性・説明可能性・追跡可能性はいわゆる**説明可能な AI** の構想と、安全性・堅牢性・制御可能性は製造者に対する設計・構造上の義務と、適正な利用は製造者の指示上の義務及びエンドユーザーを除く利用者に課せられる義務と[520]、モニタリング・監査は製造者の製造監視義務と、それぞれ通底する。この点、第 2 章や第 3 章第 1 節で言及した開発製造者の刑事過失責任を論じるにあたって、その根拠となる注意義務違反を認定する際の注意義務の内容は、現行の製造物責任法や車両運送法 68 条 9 項（車両のリコール義務）のような法律上の義務のみならず、法的期待状況から導くことも許されうる。そう考えると、本ガイドラインは AI 製品の開発製造者に課せられる義務内容をより具体化するものとすれば、この行政法規たるガイドラインに記載される原則をもって義務を確定させることができるともいえるだろう。しかし、これら原則を遵守するための実効性を伴う取組[521]が法的義務という形で存在するのであれば、仮に何らかの（刑）法的問題が生じた場合でもその法的根拠や意味内容の解釈で開発製造者のみならず、エンドユーザーを除く利用者を困惑させることは少なくなるだろう。これこそが法的観点から見た「安心・安全で信

---

516 総務省 AI ネットワーク社会推進会議「報告書 2022 ～『安心・安全で信頼性のある AI の社会実装』の更なる推進」～」（2022 年）1 頁以下参照。

517 総務省・前掲（注 516）48 頁では、1. 人間中心、2. 人間の尊厳、3. 多様性・包摂、4. 持続可能な社会、5. 国際協力、6. 適正な利用、7. 教育・リテラシー、8. 人間の判断の介在・制御可能性、9. 適正な習（学習データの質）、10. AI 間の連携、11. 安全性、12. セキュリティ、13. プライバシー、14. 公平性 15. 透明性・説明可能性、16. アカウンタビリティ、17. 堅牢性、18. 責任、19. 追跡可能性、20. モニタリング・監査、21. ガバナンス、22. その他（コスト、効果測定）である。

518 ここでの意味は、総務省・前掲（注 548）13 頁（別冊 1）によると、他国の AI ガイドラインを参照しつつ「ステークホルダには AI が適切な条件下で、適切な訓練を受けた人々によって使用されることを保証すること」や、「AI に基づく意思決定が、誰の健康や安全にも脅威を与えないこと」を挙げており、AI 技術によって問題が発生した際のステークホルダのアカウンタビリティ（説明責任）とは区別される。

519 総務省・前掲（注 516）13 頁（別冊 1）によると、透明性を保障する意味で「AI に基づく意思決定に影響を与えたデータや意思決定の根拠を追跡できること」を根底に置くと考えられる。

520 総務省・前掲（注 516）7 頁（別冊 1）によると、他国の AI ガイドラインを参照しつつ、「個人は、どのようなデータが収集され、何のために、どのような状況で使用されるかについて自分自身でコントールできるようにするべきである」や「人間は、意思決定と行動を自律的システムに委ねるかどうか、いつ、どのように委ねるかを選択し、適正に利用しなければならない」ということを例示している。これらは日本「AI 開発ガイドライン」・「AI 利活用ガイドライン」では言及されていないエンドユーザーの利用者までを視野に入れたものであることに留意しなければならない。

521 総務省・前掲（注 516）62 頁によると、事業者自身による取組のみならず、リスクを洗い出すフレームワークの構築や外部機関によるモニタリングの仕組みの整備、チェックシートの策定や認定制度の創設があげられている。

頼性のある AI の社会実装」[522]を実現する一助となるし、この見地は中国の「倫理規範」が大いに参考となるだろう。

第 3 款　小括

本節で見た 2010 年代後半から 2020 年代初頭にかけての国内外の AI 開発や利活用に関する法規則・ガイドラインでは、強い規制を伴うものや法的拘束力のない諸原則にとどまるものなど様々な性格を有するものがあった。特に開発規制の観点から見れば、将来に対する規制を意味することになるため、欧州 AI 規制案のようにその原則で直接に制裁（罰則）的規制を設けるのは、技術開発・販売流通を委縮させてしまう可能性を秘めている以上首肯しかねるところである。制裁や罰則はその規制の実効性を担保するための最終手段であるから、可能な限り最小限度の事例に留めるべきであり、むしろ実効性を担保するならば**説明可能な AI** の構築や外部機関による監視制度、行政機関による認定制度の創設を先行すべきであると思われる。ただし付言すると、米国や日本のように、もっぱら AI 製品の製造者やエンドユーザーを除く利用者のようなステークホルダを念頭に置いた法規になっていることに留意しなければならない。AI 製品に関与する主体にはステークホルダのみならずエンドユーザーたる利用者も含まれるし、この利用者に対する適正な利用を求めることも忘れてはならない。そこで参考となるのが中国の倫理規範であり、ステークホルダとエンドユーザーすべてを包含する法規を策定することこそが重要である。その一方で、法律によって定められる義務となることがその実効性を確実なものにすることを可能にし、「安心・安全で信頼性のある AI の社会実装」を実現することができるだろう。

第 3 節　刑法上の保護

「AI のための刑罰」という表現が持ちうる意味は、AI を構築し、それを利用する人々に対処しなければならないという示唆を含む。例えば、選挙の不正操作や道路交通違反を引き起こす可能性のある AI の利用について当てはまる。すなわち、ロボットもしくは実体を持たない AI を道具として利用する人間は、故意または過失によって惹起された刑法上の結果について、犯罪への関与者として刑事責任を負う可能性があることはこれまでの検討から明らかである。例えば、自動運転車の利用者、所有者もしくは製造者について、原則として、定められた注意基準を顧慮しなければ注意義務違反を認定することはできる。この場合、関与者にいかなる注意基準を課したいのか、また、帰責主体がどのようなものになるのかということが差し迫った問題となっていた。

例えば、自動運転車の領域では、上記各主体に課せられる義務の明文化が議論されていた。これにはレベル 4 の自動運転車の公道走行について定めた 2021 年ドイツ道路交通法改正や第 2 章第 4 節第 1 款で言及した 2022 年日本道路交通法改正が鍵となる。そのポイントとして、運転者なしでも所定の運行領域を独立して運転することができること、自然人の技術監

[522] 総務省 AI ネットワーク社会推進会議のテーマである。総務省・前掲（注 516）1 頁参照。

督者を置かなければならないこと、そして様々な法益への損害が避けられない場合は、人命保護を最優先しながら、各人の法益の重要性を考慮する。生命に対して避けられない危険が生じた場合には、個人的な特徴に基づいてさらなる重み付けをしないことが明文化された。

しかし「AI の創る者のための刑罰」を将来のために検討すべきことも否定されないだろう。研究者が故意はないものの、それに対応する兆候に反して、例えば身体を侵害するような効果のあるような危険な AI を作成または公開した場合、過失の可罰性を帯びるかという疑問がある。このとき、将来的に妥当するだろう AI 研究の注意規制が尊重されない場合、予見可能性はすでに最終的に侵害する事象の詳細な予見を要しないため、予見可能性の欠如から研究者は少なくとも免れられない[523]。ただし、研究者は、自己学習技術の効果を限定させることが難しい、もしくは研究者はこの知見を考慮に入れなければならない、ということについては承知しているはずである。さらに、極めて性急な刑法の投入を主張するわけではないが[524]、しばしばボーダーラインとなる過失責任を超えて、将来のロボット、ないしは AI の単なる開発が部分的に刑法上制限されるおそれもある。しかし、法を危険にさらすような AI の作成ないしは普及の将来的な禁止は、その有害な投入もしくは刑法上の態度規範としての AI の所為に先立ち、例えば AI 兵器システムの文脈で真摯に考慮に入れられるが、そのための検討事項が 2 つ挙げられる[525]。

法を危険にさらす機械というアクターの創造から法を守るための規範は、それ自体が厳格な犯罪化の基準を満たすことができる。すべての人間の利益のために、法とそれによる最高度の個人法益の保護[526]を信頼することができるようにする。ただし、未知の AI の所為に対する研究者への処罰は、そもそも原理的に不確実なものであるため、特にその前倒しは問題となることに留意しなければならない。

次に、制裁で強化された禁止をテーマ化するのはもはや尚早なものでないと考えられる。刑事立法は、基本的かつ、時間的にはまだ切迫しない徹底した議論に基づくべきであり[527]、それに加えて現在の AI のユーフォリアの中では、考慮すべき危険性について慎重に公式化

---

523 ここでは客観的帰属に限定されないであろう広範な意識なき過失の予見可能性基準については,BGHSt 48, 34,39; BGH NStZ 2001, 143, 144 f.や本稿第 2 章第 2 款第 4 項を参照。その他に *Joerden,* a.a.O. (fn.267) S. 195, 207 ff.; *Gleß/Weigend*, (fn.153), 561, 581 f.,

524 とりわけ刑法の投入が早まることへの批判については,例えば包括的な *Puschke,* Legitimation, Grenzen und Dogmatik von Vorbereitungstatbeständen, S. 49 ff., 137 ff. 現在でも、刑法 30 条、159 条の例外を除き、実用性に乏しい準備罪がいくつかある。ドイツ刑法の領域では、例えば、コンピュータ刑法では 202c 条、テロリズムの分野では、89 条 a、89 条 b、財産刑法における 265 条、医事刑法における 217 条が、関連する問題として 184 条 i がある(*Gaede* a.a.O.(fn. 472), S.82 Fn.280)。日本刑法では、組織的な犯罪の処罰及び処罰収益の規制等に関する法律 6 条の 2 がこれに該当する。

525 *Gaede*, a.a.O.(fn.472), S.82 ff.

526 特別な生命と健康の意味につき、たとえば注意義務の決定の文脈においては *Gless/Weigend*, a.a.O.(fn.153) S. 561, 584 f.も参照。

527 この文脈ではすでに *Sehuhr,* Willensfreiheit, Roboter und Auswahlaxiom, in: *Hilgendorf/Beck*,a.a.O.(fn.8) , S. 43 f.; *Simmler/Markwalder*, Roboter in der Verantwortung?, ZStW 129 (2017), S.20, 22.で指摘がある。

された際には、単なる象徴的に留まらない法的根拠を早い段階で設定する必要がある[528]。今後数十年の間に危険な AI の萌芽が出現するか否かは断言できないものの、世界規模の AI 研究の開示義務を負わない状態をあらゆる国家や社会システムを超えて細部まで看破することは困難である。また、ある一国の意見は世界のすべての研究所に到底届くわけではないので、規制は早い段階で始めなければならず、それには、AI 研究の有害な継続に対する世界規模での規制が理想的である[529]。そうでなければ、例えば AI 兵器が一度本格的に利用されうるようになれば、それを持たざる国家は虚しくも国際的に AI 兵器の排斥を要求することになるだろう。

ただし、私見としては、実例の乏しいものに対する「抽象的な可能性」の枠組での規制には慎重になるべきであると考える。なぜなら、これでは具体的危険のみならず、抽象的な危険で制裁が、ともすれば刑法の枠組にける刑罰が発動されることになってしまうからである。そのため、立法上で AI 研究も含めて、一定の基準を設け、たとえば、許可義務・届出義務・免許制度などを仔細に設定して、各々の主体に課される義務を明確化すること[530]により、従前よりも望ましい展開を期待することができるのではないだろうか。もちろん、AI は非常に多彩かつ多岐にわたるものであるから、この文脈で刑法が実際にどこまで進むべきか、意味のある効果を発揮するためにそのような規範を技術に配慮した形でどのように策定すべきか、国際的にどの程度まで法を実現できるのかは、今後の法研究や政策的な助言によってまとめ挙げられるべきものであから、今後の AI 規制に係る国内外の動向には常に注視しなければならない。

528 *Gaede*, a.a.O. (fn.472), S.84
529 *Gaede*, a.a.O. (fn.472), S.84
530 総務省・前掲「AI 開発ガイドライン」（注 511）や「AI 利活用ガイドライン」（注 512）では求められる諸原則やその実効性を担保する法創設が提案されているものの、それがステークホルダに限定されており、エンドユーザーにまで及ばないことは前節で指摘したとおりである。

おわりに

本論文では、AI の利活用における刑法上の諸問題というテーマで、主に AI 製品に関与する主体である製造者と利用者を中心に、AI 製品が人間の生命・身体・財産を侵害した場合、さらには経済犯罪を遂行した場合や AI 製品がサイバー攻撃を受けた際の刑法上の評価について検討を行った。

その際、特に先行研究の蓄積がある AI を搭載した自動運転車の事故に伴う人間の生命・身体を侵害した事例での検討において論者によって想定する AI の定義が確定していなかったことを端緒に、その刑法上の問題を論じる前にまずは AI の研究史にさかのぼって AI の定義を確定させることを試みた。しかし、その当時から AI の定義には困難を伴っていたことから、現存する AI の現象形態から帰納する形で、特定のタスクの遂行に特化し自律的判断する能力を有しないいわゆる「弱い AI」をこれら刑法上の問題を論じる上で対象とすべきであることが確認された（第 1 章）。

その上で、第 2 章では本論文の主題でもある AI 製品の事故に伴う人間の生命・身体の侵害事例を検討した。そこでは、道交法上の義務が創設された自動運転車の事故事例と、いまだ法律上の義務の存在しない、例えば介護用ロボットや産業用ロボットの利用中の事故事例とに分類して検討を行った。自動運転車の場合は、レベル 4 ではその自動運転車の運行に関与する主体の義務が仔細に規定されている一方で、レベル 3 については画面注視に係る義務と点検整備に係る義務のみが、さらにレベル 2 以下では普通自動車の運転手と同様の義務内容が条文の解釈上課せられる。しかし、レベル 2 の自動運転車の事故に関する判例の検討スキームにも見られるように、およそ普通自動車の操作と同一とはいえないレベル 2 の自動運転車にも普通自動車と同様の運転手の義務が課せられるとするのは不適切だろう。この点については、レベルに即した運転手の義務付けが必要になると思われる。

それとは対照的に、法律上の義務の存在しない AI 製品の事故事例に関しても、差し当たっては自動運転車の事例と同様に、その AI 製品に関与する主体たる人間に刑事責任が帰属しうると考えるべきである。2010 年代から盛んに議論されてきたこのテーマでは、事故に至る動作をする判断をしたのが人間の判断ではなく AI の学習による自立的判断であることや、AI の学習経過がブラックボックス化して人間の判断と事故結果の間の因果関係が遮断される可能性などを考慮すると帰属主体が存在しなくなることが指摘されてきた。これに対し、AI 自体に刑事責任を帰属させようとする試みや製造者に対する厳格責任モデルを用いようとする試み、不規制によって刑事責任の帰属を考慮しないとする試みなどが考察されたが、いずれも伝統的解釈から外れたものとなりうるし、新たな刑罰モデルの構築を必要とするが、AI 自体への刑事責任帰属では、本来であれば刑事責任が帰属されるべき主体が AI を隠れ蓑にして責任帰属から逃れる可能性が否定できなかったり、厳格責任モデルでは製造者に過度な負担を課したりすることになりかねず、AI 製品に関与する主体間でアンバランスな解決に至ってしまう。そこで、自動運転車のように、他の AI 製品に関しても一定

の法的義務に基づいてAI製品の関与者の義務を確定すべきである。例えば製造者に対しては、製造物責任法下で課せられる製造上の義務、設計上の義務、指示・警告上の義務（製品監視義務）を手掛かりに、刑法上の製造物責任を検討すべきである。ここで留意すべきは、製造者はこれら義務に違反したからといって直ちに刑法上の過失を構成するのではなく、その義務の内容の保護目的に従って、生じた結果との因果関係の有無を慎重に検討することである。製造者側で関与する技術サービスプロバイダや許可責任者たる国家・地方公共団体に対しても、当該AI製品の供給・流通に関する一定の明文の義務付けをして、その義務内容の保護目的に従って、生じた結果との因果関係の有無を検討するというスキームが望ましい。その一方で、利用者や所有者に対しては製造者側の指示を遵守し、これを悪用・濫用しないようにするという一定の義務付けも必要とされるだろう。多くのAIガイドラインや法規ではこの観点があまり考慮されていないが、実際にAI製品を使用するのは利用者や所有者といったエンドユーザーなので、これら主体に対してこうした義務付けを製造者側と並置する形で明文化することもまた重要である。もちろん、結果帰属の検討の際にはその義務内容の保護目的に従って、生じた結果との因果関係の有無を検討することは上記と同様である。

しかし、過失犯処罰規定のない経済犯罪の場合は上記スキームで検討するには困難を伴う。そこで、第3章第2節では実体を持たないAI製品であるAI・アルゴリズムが、その学習の結果、利用者の知らないところで相場操縦、インサイダー取引のような証券犯罪、価格の協調的行為のような競争法違反を遂行した場合について検討した。証券犯罪では、まずその刑罰の根拠となる金融商品取引法上の解釈が問題となる。相場操縦では、たとえ利用者の知らないところで相場操縦的取引が遂行されたとしても、客観的事実に基づいて利用者の故意が推定されるといった実務上の運用があるため、相場操縦行為が認定される可能性が否定できない。そうすると、利用者は常に相場操縦の可能性を考慮しながらAI・アルゴリズムを利用することになるが、これでは利用上も法的にも過度な負担になりうるし、開発普及を阻害する結果になりかねない。だからこそ、相場操縦でないと認められるに足るシステム構築が製造者には求められ、その判断プロセスを明確にする構造―説明可能なAI―が重要である。また、インサイダー取引では未公開重要事実を知った利用者がその事実を利用したか否かは問わず証券等の取引が行われ、もって利用者に一定の利益がもたらされた場合には形式的にインサイダー取引に該当するが、適用除外要件に該当する限りでインサイダー取引は成立しないという構造となっている。AI・アルゴリズムの利用促進、開発普及の観点から、この場合も適用除外要件に適合する形でのシステム構築が製造者には求められ、そこでもその判断プロセスを明確にする構造―説明可能なAI―が重要である。それに対して、学習するAI・アルゴリズムによりその利用者間（競争事業者間）での価格協調が実現した場合は、独占禁止法の解釈が関わってくる。判例上は黙示による競争事業者間の「共同性」も認定していることから、たとえ上記の場合でも不当な取引制限に該当しうるが、そもそもAI・アルゴリズムの学習により価格協調行為が遂行される可能性は現状では低いことや、

「不当な取引制限」に対しては排除措置命令、課徴金、刑事罰の可能性が規定されていることを考慮すれば、「共同性」要件の認定には特に慎重になるべきであるし、「不当な取引制限」に該当するとすべきでもない。この経済犯罪における先行研究で目立っていたのは、刑事罰の可能性が予定される類型であるにもかかわらず、利用者の処罰を制限する論調ではなかったことである。このことは AI や AI 学習特有の問題ではないが、刑事罰を発動する可能性を秘める以上、その構成要件の認定を慎重する姿勢が必要であると考える。

AI 製品が行為客体になる場合、すなわち AI 製品がサイバー攻撃を受けて利用者の情報が取得されたり、その内部データの変更・破壊により製品利用を妨げられたりした際の刑法上の評価では、構成要件の問題と AI の学習のブラックボックス性の問題に大別される。

構成要件の問題では、まず AI 製品にハッキングする行為が必ずしも不正アクセスを構成しないこと、AI 製品内に記録された利用者情報の取得する行為は特別刑法上の問題であること、そして内部データの変更・破壊の場合ではその製品利用に業務性があり、それが業務妨害結果に至ったという限りで電子計算機損壊等業務妨害罪が成立するにすぎないことである。それゆえ、私的空間に属する AI 製品の内部データの変更・破壊に関しては、それがアクセス制御機能を有している限りで不正アクセス罪のみが成立するという帰結になる。さらに、考慮すべきもう一つの事例として、資産運用を行う AI ソフトウェア・エージェントに対して不正なデータが用いられ、結果として AI ソフトウェア・エージェントの利用者に対して財産的損害が発生したものである。ここでは電子計算機使用詐欺罪の成否が検討されるが、条文解釈上行為者に対して不当利得を要求するため、本罪の成立は認められず、その AI・ソフトウェア・エージェントの利用が利用者の業務に属し、財産的損害が利用者の業務を妨害したといえる限りで電子計算機損壊等業務妨害罪が成立する。ここには、電子計算機損壊等業務妨害罪の成否と同様の問題があることに注意しなければならない。

AI の学習のブラックボックス性の問題は、電子計算機損壊等業務妨害罪では、その損壊もしくは虚偽の情報または不正な指令の供与の原因が行為者によるものなのか、AI の学習によるものなのかが不明であった場合が、電子計算機使用詐欺罪では、財産権の得喪もしくは変更に係る不実の電磁的記録の原因たる虚偽の情報もしくは不正な指令、ないしは財産権の得喪もしくは変更に係る虚偽の電磁的記録が行為者によってもたらされたのか、それとも AI ソフトウェア・エージェントの学習によってもたらされたのかが不明な場合にある。両者とも、その因果関係が不明確な場合はたとえ利用者の業務が妨害されたとしても、行為者が不当利得を得た結果が生じたとしても、行為者には未遂罪が成立するにすぎないという帰結となってしまう。ここでも重要なのが説明可能な AI の構想であり、本来ならば刑事責任が帰属されるべき行為者（攻撃者）が AI の学習を理由に未遂減軽の可能性を残してしまう状況を防ぐことが実現できる。

第 4 章では、2010 年代後半から各国で進められている将来的な AI 開発の指針・規制は、2020 年代に入るとその内容に変化を見せていることを確認した。その規定ぶりは強い制裁規範を持つものから、法的拘束力を持たないガイドラインにとどまるものまで様々である

が、将来の開発に対する規制を論じる際には、いまだ具体的な危険のない状態で過度な制裁を課すことは、将来的な AI 開発を委縮させる効果を招来することに注意しなければならない。もちろん、人権を侵害するような開発に関しては規制の対象とすべきであるが、現状の技術水準を考慮すればそのような開発が表立って行われているわけではないので、直ちに強い規制を必要とするわけではない。しかし AI を搭載した製品は日々進歩を続けており、数多くの人間の主体が AI と関わるようになっているので、これら主体が遵守すべき原則、課せられる義務を具体的に作成すべき時期に差し掛かってきているように感じる。その実効性を担保するための許認可、監査制度などのソフトな措置から創設し、エンドユーザーたる利用者の利益と製造者側の負担とのバランスを考慮しながら、AI 製品を取り巻く主体が遵守すべき法律上の原則・義務を創設することが、これからの AI 研究開発、ひいては販売流通・利活用にとって不可欠なものである。

残された課題として、第 1 章や第 4 章第 2 項で言及した各国の AI に関する法規で述べられた、①AI 製品の利用者のデータ保護およびその第三者利用に関する刑法上の観点、②ディープウェブ内で AI を用いたプラットフォーム事業者の刑法上の責任、③国際的な証券取引のレベルでの経済犯罪が遂行された場合の刑法上の解釈である。①について日本では 2022 年個人情報保護法改正により罰則規定が厳罰化されたことも踏まえ、現在の喫緊の課題ともいえる。②・③については先行研究[531]が本年になって立て続けに刊行されているが、ネットワークによってグローバル化した社会の中で新たに考慮される重要な課題といえる。これらについては、別途検討を行いたいと思う。

[531] *Grimm*, Das Insiderhandelsverbot zwischen Rechtstheorie und Rechtspraxis, Nomos 2022; *Weber*, Die Strafbarkeit von Plattformbetreibern im Darknet, Nomos 2022.

参考文献

（日本語文献）

アルゴリズム・AI の利用を巡る法律問題研究会「投資判断におけるアルゴリズム・AI の利用と法的責任」金融法務(2019 年)

Asada Research Group「子供アンドロイドの開発」(http://www.er.ams.eng.osaka-u.ac./jpasadalab/?page_id=177)（最終アクセス 2022 年 11 月 28 日）

浅田和茂「ファイル共有ソフト利用者に「イカタコウィルス」を受信・実行させた行為が器物損壊罪に当たるとされた事例」新・判例解説 Watch（法学セミナー増刊）11 号

浅田和茂「判例に見られる罪刑法定主義の危機」立命館法学 345・346 号（2012 年）

浅田和茂『刑法各論（第 2 版)』(成文堂、2020 年）

安達光治「危険の現実化論について」井田良ほか編『浅田和茂先生古稀祝賀論文集（上巻)』(成文堂、2016 年）

石黒共生ヒューマンロボットインタラクションプロジェクト（https://www.jst.go.jp/erato/ishiguro/outline.html）(最終アクセス 2022 年 11 月 28 日)

泉眞樹子「ドイツにおける自動運転車の公道通行 ―第 8 次道路交通法改正―」国立国会図書館（2018 年）https://dl.ndl.go.jp/view/download/digidepo_11052071_po_02750004.pdf?contentNo=1（最終アクセス 2022 年 11 月 28 日）

市川芳治「人工知能（AI）時代の競争法に関する一試論 ～“アルゴリズム”によるカルテル：欧米の最新事例からの示唆を受けて～（下)」国際商事法務 45 巻 2 号（2017 年）

伊藤榮樹・小野慶二・荘司邦雄『注釈特別刑法 第 6 巻 交通法・通信法編 II』(立花書房、1994 年）

稲垣悠一『欠陥製品に関する刑事過失責任と不作為犯論』(専修大学出版、2014 年）

稲谷龍彦「人工知能搭載機器に関する新たな刑事法規制について」法律時報 91 巻 4 号（2019 年）

稲谷龍彦「ロボット事故の刑事責任」日本ロボット学会誌 1 巻 38 号（2020 年）

稲谷龍彦「Society 5.0 における新しいガバナンスシステムとサンクションの役割（上)」法律時報 94 巻 3 号（2021 年）

稲谷龍彦「Society 5.0 における新しいガバナンスシステムとサンクションの役割（下)」法律時報 94 巻 4 号（2021 年）

岩原紳作ほか「金融商品取引法セミナー（第 17 回）追補：内部者取引規制と公開買付規制」ジュリスト 1417 号（2011 年）

岩間康夫「刑法上の製造物責任と先行行為に基づく保障人的義務―近時のドイツにおける判例及び学説から」愛媛法学会雑誌 18 巻 4 号（1991 年）

岩間康夫「製造物責任の事例における取締役の刑事責任―集団的決定に関与した者の答責―」愛媛法学会雑誌 22 巻 1 号（1995 年）

岩間康夫『製造物責任と不作為犯』（成文堂、2010 年）

岩間康夫「刑事製造物責任の諸論点―とりわけ回収義務の根拠に関するドイツの議論について―」刑事法ジャーナル 37 号（2013 年）

伊庭斉志『ゲーム AI と深層学習　ニューロ進化と人間性』（オーム社、2018 年）

伊藤嘉亮「エリック・ヒルゲンドルフ『ロボットは有責に行為することができるか？ 規範的な基本語彙の機械への転用可能性について』（文献紹介『ロボットと法』シリーズの論文紹介(1)）」千葉大学法学論集 31 巻 2 号（2016 年）

今井猛嘉「AI 時代の刑事司法」罪と罰 222 号（2019 年）

今井猛嘉「自動車の自動運転と刑事実体法―その序論的考察」西田典之先生献呈論文集（有斐閣、2017 年）

石井徹哉「AI に関する刑法上の課題」罪と罰 222 号（2019 年）

ヴァルター・ペロン（高橋則夫訳）「刑法における製造物責任―ドイツ連邦通常裁判所「皮革用スプレー判決』をめぐって―」東洋大学比較法 31 号（1994 年）

遠藤聡太「人工知能(AI)搭載機器の安全性確保義務と社会的便益の考慮」法律時報 91 巻 4 号（2019 年）

大塚仁・河上和雄・中山善房・古田佑紀『大コンメンタール刑法（第 3 版)』（青林書院、2018 年）

大塚仁『刑法概説 総論（第 4 版)』（有斐閣、2008 年）

大森泰人「課徴金(下)」金法 1896 号（2010 年）

大谷實『刑法講義各論（第 5 版)』（成文堂、2019 年）

岡部雅人「刑事製造物責任における『回収義務』について」早稲田大学大学院法研論集 123 号（2007 年）

岡部雅人「刑事製造物責任における回収義務の発生根拠―わが国の議論状況をめぐって一」刑事法ジャーナル 37 号（2013 年）

外務省「8 ヵ国デンヴァー・サミット コミュニケ（仮訳)」（1997 年)。

外務省「8 カ国司法・内務閣僚級会合 1997 年 12 月 9-10 日　コミュニケ（仮訳)」（1998 年）

甲斐勝則「欠陥製品の製造・販売と刑事過失」斉藤豊治・日高義博・甲斐勝則・大塚裕史編『神山敏夫先生古稀祝賀論文集』（成文堂、2006 年）

角田正紀「判批」判評 356 号（1988 年）

葛西まゆこ「イントロダクション 憲法学から見た通信の秘密」警察学論集 66 巻 2 号（2013 年）

金井貴嗣・川濵昇・泉水文雄編『独占禁止法（第 6 版)』（弘文堂、2021 年）

川口浩一「ロボットの刑事責任 2.0」刑事法ジャーナル 57 号（2018 年）

川口浩一「ロボット・AI に対する刑罰をめぐる最近の議論」法律論叢 94 巻 4・5 号（2022 年）

神田秀樹・黒沼悦郎・松尾直彦編『金融商品取引法コンメンタール 4 不公正取引規制・課徴金・罰則』（商事法務、2011 年）

金山博・武田浩一「Watson：クイズ番組に挑戦する質問応答システム」情報処理 52 巻 7 号（2011 年）

貴島逸斗「巨額制裁金が現実に 『AI 倫理』待ったなし」日経コンピュータ 1075 号（2022 年）

北川佳世子「製造物責任をめぐる刑法上の問題点―ドイツ連邦通常裁判所の皮革用スプレー判決をめぐる議論を手掛かりに―」早稲田法学 71 巻 2 号（1996 年）

北和樹「EU が目指す AI 社会のための規制法」立命館大学人文科学研究所紀要 131 号（2022 年）

木村亀二「不作為犯における作為義務」同『刑法解釈の諸問題第一巻』（有斐閣、1939 年）

木目田裕監修・直村あさひ法律事務所・危機管理グループ編『インサイダー取引規制の実務（第 2 版)』（商事法務、2014 年）

金融審議会市場制度ワーキング・グループ「最良執行のあり方等に関するタスクフォース 報告書」（2021 年）

金融財政事情研究会「対岸の火事ではない欧州の AI 規制案 : 新聞の盲点」金融財政事情 73 巻 4 号（2022 年）

金融商品取引法研究会「インサイダー取引規制と自己株式」（2015 年）

金融庁 証券取引等監視委員会「インサイダー取引規制に関するＱ＆Ａ」（2019 年）

久木田水生「AI のリスクと倫理」第 36 回人工知能学会全国大会論文集（2022 年）

栗原佑介「カナダ、欧米における AI 規制法案の動向からみる AI ガバナンス」InfoCom T&S world trend report 401 号（2022 年）

黒沼悦郎『金融商品取引法[第 2 版]』（有斐閣, 2020 年）

ゲーツェル・ベン「汎用人工知能概観」人工知能 29 巻 3 号（2014 年）

経済産業省「GOVERNANCE INNOVATION Ver.2: アジャイル・ガバナンスのデザインと実装に向けて」（2021 年）

小泉雄介「欧州 AI 規制案の概要」データ社会推進協議会データ倫理プライバシー研究 WG 資料（2021 年）https://www.i-ise.com/jp/information/report/2021/202106.pdf（最終アクセス 2022 年 11 月 28 日）

厚生労働省「介護ロボット導入活用事例集 2021」（2021 年）
https://www.mhlw.go.jp/content/12300000/000928395.pdf（最終アクセス 2022 年 11 月 24 日）

厚生労働省ホームページ https://www.mhlw.go.jp/file/06-Seisakujouhou-12300000-Roukenkyoku/0000210895.pdf（最終アクセス 2022 年 11 月 24 日）

小林一郎『人工知能の基礎』（サイエンス社、2008 年）4 頁。

斉藤豊治・浅田和茂・松宮孝明・髙山佳奈子『新経済刑法入門 (第 3 版)』（成文堂、2020 年）

佐伯千仞『刑法講義 総論（4 訂版)』（有斐閣、1984 年）

佐伯仁志「インサイダー取引」西田典之編『金融業務と刑事法』（有斐閣、1997 年）

佐伯仁志「保障人的地位の発生根拠について」内藤謙ほか編『香川達夫博士古稀祝賀論文 刑事法学の課題と展望』（成文堂、1996 年）

佐久間修「AI と刑法・序説」名古屋学院大学論集社会科学編 55 巻 1 号（2018 年）

佐久間修「AI の刑事責任―否定説の観点から」刑法雑誌 59 巻 2 号（2020 年）

坂下陽輔「人工知能の開発・利用における過失―自動運転車と過失を題材に」法律時報 91 巻 4 号（2019 年）

笹倉宏紀「人工知能の法規制における行政手続と刑事手続」法律時報 91 巻 4 号（2019 年）41 頁以下

塩見淳「瑕疵ある製造物を回収する義務について」刑法雑誌 42 巻 3 号（2003 年）

鎮目征樹「刑事製造物責任における不作為犯論の意義と展開」本郷法政紀要 8 号（1999 年）

鎮目征樹・西貝吉晃・北條孝佳『情報刑法I サイバーセキュリティ関連犯罪』（弘文堂、2022 年）

証券取引等監視委員会「証券取引等監視委員会の活動状況(平成 20 年度)」（2009 年）

消費者庁消費者安全課『逐条解説 製造物責任法（第 2 版)』（商事法務、2018 年）

白石忠志「独占禁止法〔第 3 版〕」（有斐閣、2016 年）

杉山徳明・吉田雅之「『情報処理の高度化等に対処するための刑法等の一部を改正する法律』について」曹時 64 巻 4 号（2012 年）

泉水文雄『独占禁止法』（有斐閣、2022 年）

総務省 AI ネットワーク社会推進会議「報告書 2022 ～『安心・安全で信頼性のある AI の社会実装』の更なる推進」～」（2022 年）

総務省「AI ネットワーク社会推進会議 報告書 2021～『安心・安全で信頼性のある AI の社会実装」の推進～』（2021 年）

総務省「AI 利活用ガイドライン～AI 利活用のためのプラクティカルリファレンス～」（2019 年）

総務省「国際的な議論のための AI 開発ガイドライン案」（2017 年）

総務省「令和元年度版 情報通信白書 第 1 部 第 3 節 2 .AI に関する動向 (1)」（2019 年）

園田寿「『イカタコ事件』について」：器物損壊罪における「損壊」の概念〈判例批評〉甲南法務研究 8 号

大証金融商品取引法研究会報告「市場監視の実際」（2010 年）

多賀谷一照監修 電気通信事業法研究会編著『電気通信事業法逐条解説 改訂版』（2019 年、一般財団法人情報通信振興会）

武田邦宣「不当な取引制限における意思の連絡要件」日本経済法学会年報 37 号（2016 年）

谷口忠大『イラストで学ぶ 人工知能概論 [改訂第 2 版]』（講談社、2020 年）205 頁。

団藤重光『刑法綱要総論（第 3 版）』（弘文社、1990 年）

田村翔「サシャ・ツィーマン『機械の本性とは何であったか？ 機械刑法をめぐる議論について』（文献紹介『ロボットと法』シリーズの論文紹介(2)）」千葉大学法学論集 31 巻 3 号（2016 年）

張小寧「証券犯罪の総合的研究（2）―実効的規制のための基礎的考察―」立命館法学 343 号（2012 年）

土倉澄子『逐条講義 製造物責任法 第 2 版 基本的考え方と裁判例』（勁草書房、2018 年）

デジタル市場における競争政策に関する研究会「アルゴリズム／AI と競争政策」（2021 年）

寺田麻佑・板倉陽一郎「欧州（EU）における 2021 年 AI 規制法案をめぐる各種意見と EU の対応の検討」情報処理学会研究報告 22 号（2022 年）

道路交通執務研究会編著（野下文雄原著）『執務資料 道路交通法解説〔18 訂版〕』（東京法令出版、2020 年）

鳥海不二夫「人工知能技術を俯瞰する」立法と調査 405 号（2018 年 10 月）

中川由賀「具体的事故事例分析を通じた自動運転車の交通事故に関する刑事責任の研究② ～運転支援車 (レベル２) の事故～」中京法学 55 号 1 巻（2021 年）

中山研一『刑法総論』（成文堂、1982 年）

西貝吉晃「コネクティッドカーシステムに対するサイバー攻撃と犯罪」法律時報 91 巻 4 号（2019 年）

西貝吉晃「コンピュータ・サボタージュ罪 刑法 303 条 b」刑事法ジャーナル 71 号（2022 年）

西田典之（橋爪隆補訂）『刑法総論（第 3 版）』（弘文堂、2019 年）

西田典之・山口厚・佐伯仁志編『注釈刑法 第 2 巻 各論（１）』（有斐閣、2020 年）

西田典之・山口厚・佐伯仁志編『注釈刑法 第 4 巻 各論（３）』（有斐閣、2020 年）

西原春夫『交通事故と信頼の原則』（成文堂、1973 年）

日本経済新聞「金融庁、アルゴリズム取引悪用の相場操縦で課徴金命令」（2011 年 2 月 16 日）

日本経済新聞「高速取引(HFT)とは　データ基に 1 秒で数千回の売買注文」（2019 年 10 月 20 日）

日本経済新聞「囲碁 AI、プロに 4 勝 1 敗　最終局も熱戦制す」（2016 年 3 月 15 日）

日経 EXTEC「Google 社の『Waymo』が自動運転開発に与えるインパクト」（2016 年 12 月 26 日）

根津洸希「スザンネ・ベック『インテリジェント・エージェントと刑法 過失、答責分配、電子的人格』」千葉大学法学論集 31 巻 3・4 号（2017 年）

根津洸希「ロボットの処罰可能性を巡る議論の現状について」比較法雑誌 51 巻第 2 号（2018 年）

根津洸希「ロボット・AI に対して『刑罰』を科すことは可能か」法学新報 125 巻 11 号（2019 年）

服部秀一『インサイダー取引規制のすべて』（商事法務研究会、2001 年）

樋笠尭士「自動運転（レベル２及び３）をめぐる刑事実務上の争点―レベル２東名事故を手がかりに―」捜査研究 847 号（2021 年）

樋笠尭士「自動運転レベル 4 における刑事実務―道路交通法改正案の分析と提案―」捜査研究 858 号（2022 年）

日高義博『不真正不作為犯の理論（第 2 版)』（慶應通信、1983 年）

福田平『全訂刑法総論』（有斐閣、2011 年）

藤吉弘亘「機械学習の進展による画像認識技術の変遷」計測と制御 58 巻 4 号（2019 年）

堀内捷三「製造物の欠陥と刑事責任一その序論的考察―」研修 546 号（1993 年）

前嶋匠「企業・組織犯罪における合議決定と帰属関係(二・完) : 因果関係と共同正犯・共同教唆」関大法学 54 巻 5 号(2005)

前田雅英編『条解刑法（第 4 版)』（弘文堂、2020 年）

前田巌「不正アクセス行為の禁止等に関する法律 8 条 1 号の罪と私電磁的記録不正作出罪との罪数関係」曹時 62 巻 10 号（2008 年）

松尾直彦『金融商品取引法 [第 2 版]』（商事法務、2020 年）

松尾剛行「自動運転車と刑事責任に関する考察 ロボット法を見据えて」早稲田大学大学院法務研究科臨床法学研究会 Law and practice 11 巻（2017 年）

松尾豊『人工知能は人間を超えるか―ディープラーニングの先にあるもの』（角川 EPUB 選書、2015 年）

松原仁・伊藤毅史「AlphaGo の技術と対戦」人工知能 31 巻 3 号（2016 年）

松原仁「コンピュータ囲碁の進歩」日本ロボット学会誌 35 巻 3 号（2017 年）

松宮孝明『過失犯論の現代的課題』（成文堂、2004 年）

松宮孝明「判批」立命館法学 343 号(2012 年）

松宮孝明「薬害エイズ事件厚生省ルート事件最高裁決定」医事法学 24 号（2009 年）

松宮孝明『刑法各論講義 (第 5 版)』（成文堂、2020 年）

松宮孝明『先端刑法総論』（日本評論社、2020 年）

松宮孝明『刑法総論講義(補訂第 5 版)』（成文堂、2020 年）

松宮孝明「自動運転と法」学術の動向（2020 年 5 月）

松宮孝明「自動運転をめぐる刑事法的諸問題」立命館法学 395 号（2021 年）

水田孝信「人工知能は相場操縦という不正な取引を勝手に行うか？―遺伝的アルゴリズムが人工市場シミュレーションで学習する場合―」第 34 回人工知能学会全国大会論文集（2020 年）

森住信人「イカ／タコウィルス事件：ソフトウェアの改変と器物損壊罪の成否〈刑事裁判例批評 268〉」刑事法ジャーナル 41 号（2014 年）

安富潔「情報化社会における刑事立法の役割―コンピュータ犯罪からサイバー犯罪へ―」慶応法学 42 号（2019 年）

山口厚『刑法総論(第 2 版)』（有斐閣、2007 年）

山口厚『刑法各論(第 2 版)』（有斐閣、2012 年）

山口厚『経済刑法』（商事法務、2012 年）

山中敬一「刑事製造物責任論における作為義務の根拠」関大法学 60 巻 5 号（2011 年）

山中敬一『刑法総論（第 3 版)』（成文堂、2015 年）

山下祐樹「AI・ロボットによる事故の責任の所在について」ノモス 45 巻（2019 年)

深町晋也「ロボット・AI と刑事責任」弥永真夫・宍戸常寿編『ロボット・AI と法』（有斐閣、2018 年）

横畠裕介『逐条解説インサイダー取引規制と罰則』（商事法務研究会、1989 年)

米澤慶治『刑法一部改正法の解説』（立花書房、1989 年）

劉憲権（孫文訳,松宮孝明監訳）「人工知能時代における刑事責任の変遷」立命館法学 396 号（2021 年）

ローレンス・D・バーンズ・クリストファー・シュルガン（児島修 訳）『AUTONOMY 自動運転の開発と未来』（辰巳出版、2020 年）

（外国語文献）

Arthur Kaufmann, Die ontlogische Struktur der Handlung, Skizze einer personalen Handlungslehre, in: FS-H. Mayer, 1966.

BCS, New EU AI regulations demand a 'fully professionalised tech industry' - institute for IT. 2021, Apr 22.

BT-Drs. 10/318

BT-Drs. 10/4728

BT-Drs. 16/3656

BT-Drs. 19/5880

Barton, Multimedia-Strafrecht Ein Handbuch für die Praxis,Hermann Luchterhand Verlag 1999.

Beck, Die Diffusion strafrechtlicher Verantwortlichkeit durch Digitalisierung und Lernende Systeme, ZIS 03/2020.

Beck, Intelligente Agenten und Strafrecht. Fahrlässigkeit, Verantwortungsverteilung, elektronische Personalität, Studien zum deutschen und türkischen Strafrecht - Delikte gegen Persönlichkeitsrechte im türkischen-deutschen Rechtsvergleich (Band 4), Ankara 2015.

Beulke/Bachmann, Die „Lederspray-Entscheidung“ -BGH St 37, 106. JuS 1992.

Bostrom, Superintelligence - Paths, Dangers, Strategies, Oxford University 2014.

Bottke, Krankmachende Bauprodukte-Produkthaftung aus zivil- und strafrechtlicher Sicht unter besonderer Berücksichtigung krankmachender Gebaude (Sick Building Syndrom)Teil 2. ZfBR 1991.

Brammsen, Strafrechtliche Ruckrufpflichten bei fehlerhaften Produkten?, GA 1993.

Calo, People Can Be So Fake: A New Dimension to Privacy and Technology Scholarship, Penn State Law Review 2010

Calvano et al., Artificial Intelligence, Algorithmic Pricining, and Collusion, American Economic Review, 3267,2020

Car to Car Communication Consortium, C2C-CC Manifesto, 2007.

Cirener/Radtke/Saan/Rönnau/Schluckebier, Leipziger Kommentar Strafgesetzbuch: StGB, 13. Aufl. De Gruyer.

De la Autorité councurrence, Budeskartellamt, Algorithms and Competition, 2019.

ECJ, Musique Diffusion française and Others v Commission, Judgment of 07.06.83, Joined Cases 100/80 to 103/80, para. 97

Eberl, Smarte Maschinen: Wie Kuenstliche Intelligenz unser Leben veraendert, HANSER 2016.

Engisch, Vom Weltbild des Juristen, 2. Aufl., Heidelberg 1965.

Erb/Schäfer (Hrsg.), Münchener Kommentar zum Strafgesetzbuch: StGB, 4.Aufl.

Ernst, Das neue Computersstrafrecht, NJW 2007.

Fetah-Moghadam, Innovationsverantwortung im Strafrecht: Zwischen strict liability, Fahrlässigkeit und erlaubtem Risiko – Zugleich ein Beitrag zur Digitalisierung des Strafrechts, ZStW 2018

Fischer, Strafrechtgesetzbuch, 66.Aufl., C.H.Beck 2019.

Foeste/Kreifels/Mühlbauer/Schütze/Weide/Westphalen/Wilde/Winkelbauer (Hrsg.), Produkthaftungshandbuch, 3.Auflage, C.H.Beck 2012.

Fraunhofer IAIS, Maschinelles Lernen „ on the edge “, 2019.

Gaede, Künstliche Intelligenz -Rechte und Strafen für Roboter? Plädoyer für eine Regulierung künstlicher Intelligenz jenseits ihrer reinen Anwendung, Robotik und Recht 18, Nomos 2018.

Gless/Weigend, Intelligente Agenten und das Strafrecht, ZStW 2014.

Gleß/Silverman/Weigend, If Robots Cause Harm, Who Is To Blame? Self-Driving Cars And Criminal Liability, New Criminal Law Review, Vol.19, Number 3, 2016.

Grimm, Das Insiderhandelsverbot zwischen Rechtstheorie und Rechtspraxis, Robotik und Recht 27, Nomos 2022.

Gubrud, “Nanotechnology and International Security”, Fifth Foresight Conference on Molecular Nanotechnology, November 1997.

Günther, Roboter und rechtliche Verantwortung, Herbert Utz Verlag 2016.

Günther/Münch, Legal lssues of Making a Robot “Readable”, in: Workshop on Robot Feedback in Human-Robot Interaction: How to Make a Robot,,Readable" for a Human lnteraction Partner, 21 st IEEE International Symposium on Robot and Human lnteractive Communication, 2012.

Goodfellow/Bengio/Courville, Deep Learning, The MIT Press 2016.

High-Level Expert Group on Artificial Intelligence, Ethics Guidelines for trustworthy AI, 2017.

Haenssle u.a., Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists, Annals of Oncology, Vol. 29, Issue 8, 2018,

Hassemer, Produktverantwortung im modernen Strafrecht, 2. Aufl., C.F.Müller 1996.

Herberger, „Künstliche Intelligenz“ und Recht, NJW 2018.

Hilgendorf, Strafrechtliche Produzentenhaftung in der Risikogesellschaft, Duncker & Humblot 1993.

Hilgendorf, Scheckkartenmißbrauch und Computerbetrug OLG Düsseldorf, NStZ-RR 1998, 137 JuS 1999.

Hilgendorf, Können Roboter schuldhaft handeln?, in: Hilgendorf/Beck (Hrsg.), Jenseits von der Maschine, Robotik und Recht 1, Nomos 2012.

Hilgendorf (Hrsg.), Aktuelle Herausforderungen des chinesischen und deutschen Strafrechts, Mohr Siebeck 2015.

Hilgendorf, Automatisiertes Fahren und Strafrecht – der „Aschaffenburger Fall“ , DRiZ 2018.

Hilgendorf, Autonome Systeme, künstliche Intelligenz und Roboter Eine Orientierung aus strafrechtlicher Perspektive, in:FS Fischer, 2019.

Hilgendorf, Das neue Computerstrafrecht, in: Hilgendorf (Hrsg.), Dimensionen des IT-Rechts (Das Strafrecht vor neuen Herausforderungen), Logos 2008.

Hilgendorf/Valerius, Computer- und lnternetstrafrecht, Springer 2012.

Hoyer, Die traditionelle Strafrechtsdogmatik vor neuen Herausforderungen: Probleme der strafrechtlichen Produkthaltung. GA 1996.

Meier, Verbraucherschutz durch Strafrecht? Überlegungen zur strafrechtlichen Produkthaftung nach der "Lederspray"-Entscheidung des BGH, NJW 1992.

Jakobs, Die Ingerenz in der Rechtsprechung des Bundesgerichtshofs. in: Roxin/Widmaier (Hrsg.), 50 Jahre Bundesgerichtshof-Festgabe aus der Wissenschaft Bd. IV. 2000.

Janka/Uhsler, Antitrust 4.0, European Competition Law Review, 2018.

Jeschke, Auf dem Weg zu einer „neuen KI“: Verteilte intelligente Systeme, in: Jeschke/Isenhardt/Hees/Henning

(Hrsg.), Automation, Communication and Cybernetics in Science and Enigneering 2015/2016.

Joerden, Strafrechtliche Perspektiven der Robotik, *Hilgendorf/Günther* (Hrsg.), Robotik und Gesetzgebung, Robotik und Recht 2, Nomos 2013

Searle, MINDS, BRAINS, AND PROGRAMMS, University of California, 1980

Kaplan, Artificial Intelligence: What Everyone Needs to Know, Oxford, 2016

Kuhlen, Haftung für Sorgfaltspflichtverletzungen in Unternehmen bei der Produktion von Gütern, in: *Hilgendorf* (Hrsg.) Aktuelle Herausforderungen des chinesischen und deutschen Strafrechts, Mohr Siebeck 2015

Kuhlen, Strafhaftung bei unterlassenem Rückruf gesundheitsgefahrender Produkte - Zugleich Anmerkung zum Urteil des BGH vom 6. 7. 1990-2 StR 549/89 (NStZ 1990. 588).

Kuhlen, Strafhaftung bei unterlassenem Rückruf gesundheitsgefährdender Produktverantwortung, NJW 1990.

Lenzen, Künstliche Intelligenz Was sie kann & was uns erwartet? , C.H. Beck 2018.

Lima, Could AI Agents Be Held Criminally Liable? Artificial Intelligence and The Challenges For Criminal Law, South Carolina Law Review 69, 677, 2018

Liptak, Sent to Prison by a Software Program's Secret Algorithms, New York Times(May 1st, 2017).

Lohmann, Strafrecht im Zeitalter von künstliche Intelligenz, Robotik und Recht 24, Nomos 2021.

Leupold/Glossner, Münchener Anwaltshandbuch IT-Recht, C.H.Beck 2011.

Maihofer, Der Handlungs Begriff im Verbrechen System, Mohr 1953.

Maihofer, Der soziale Handlungsbegriff, FS-Eb. Schmidt, 1961.

Marberth-Kubicki, Computer- und Internetstrafrecht, C.H. Beck 2009.

Marks, US Product Liability Law, International Business. LAWYER 2, 69, 1998.

Matthias, Automaten als Träger von Rechten, Logos 2008.

McCarthy, Minsky, Rochester, Shannon, A proposal for the Dartmouth summer research project on artificial intelligence (Aug. 31st, 1955).

McCarthy, WHAT IS ARTIFICIAL INTELLIGENCE? , Computer Science Department, Stanford University, 2003.

Mercedes Benz, Car-to-X Communication. Mercedes-Benz is starting a Europe-wide cooperation project.

Minoru Asada, Towards Artificial Empathy. International Journal of Social Robotics, Vol.7, No.1, 2015

Moliroris/Klindt, Produkthaftung und Produktssicherheit – Ein aktueller Rechtsprechungsüberblick, NJW 2008.

OSTP, Blueprint for an AI Bill of Rights: A Vision for Protecting Our Civil Rights in the Algorithmic Age, October 14th, 2022.

Owen, Inherent Product Hazards, Kentucky Law Journal: Vol. 93: Issue. 2, Article 4

Puschke, Legitimation, Grenzen und Dogmatik von Vorbereitungstatbeständen, Mohr Siebek 2017.

Quarck, Zur Strafbarkeit von e-Personen, ZIS 04/2020.

Rich, Artifical lntelligence, McGraw-Hill, New York 1983.

Roxin, Strafrecht Allgemeiner Teil, Band II, 3. Aufl., C.H. Beck 2003.

Roxin/Greco, Strafrecht Allgemeiner Teil, Band I, 6. Aufl., C.H.Beck 2020.

Russell/Norvig, Artificial Intelligence - A Modern Approach-, 3rd Ed., 2010.

Russell/Norvig, Artificial Intelligence -A Modern Approach-, 4th Ed., 2021.

Salaschek/Serafimova, Preissetzungsalgorithmen im Lichte von Art. 101 AEUV, Wirtschaft und Wettbewerb 2018.

Scheffer/Jonathan, One Jump Ahead: Challeging Human Spremacy in Checkers, Springer 2009

Schönke/Schröder, Strafgesetzbuch: StGB, 30. Aufl., C.H.Beck 2019

Schünemann, Unternehmenskriminarität und Strafrecht, Heymann 1979.

Schünemann, Unternehmenskriminalität, in: *Roxin/Widmaier* (Hrsg.), 50 Jahre Bundesgerichtshof-Festgabe aus der Wissenschaft Bd. IV. 2000.

Seelmann, Zurechnung zu Künstlicher Intelligenz?, FS Reinhard Merkel Teilband I, 2020.

Sehuhr, Willensfreiheit, Roboter und Auswahlaxiom, in: *Hilgendorf/Günther* (Hrsg.), Robotik und Gesetzgebung, Robotik und Recht 2, Nomos 2013.

Simmler/Markwalder, Roboter in der Verantwortung ?, ZStW 129 (2017)

Smith, Proximity-Driven Liability, Geogetown. Law Jounal 102, 2014,

Stell/Krüger, Lernen und Sicherheit in Interaktion mit Robotern aus Maschinensicht, in: Hilgendorf/Günther, Robotik und Gesetzgebung, Robotik und Recht 2, Nomos 2012

Stilwell, Warning: You May Possess Continuing Duties After the Sale of Your Product!, The Review of Litigation 26, 4, 2007.

Sutton/Barto, Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning series), Bradford Books, 1998.

Terwilleger, Navigating the Road Ahead: Florida Autonomous Vehicle Statute and Its Effect on Liability, Florida Bar Journal 89, 7, 2015.

Valerius, Strafrechtliche Grenzen lernender KünstlicherIntelligenz, GA 3/2022.

Weber, Die Strafbarkeit von Plattformbetreibern im Darknet, Robotik und Recht 25, Nomos 2022.

Welzel, Das Deutsche Strafrecht 11 Aufl., De Gruyter 1969.

Wessels/Beulke, Strafrecht Allgemeiner Teil, 52. Aufl., C.F. Müller 2022.

Wolter/Hoyer, SK-StGB-Kommentar, 6 Bd., Carl Heymanns 2022.

von Liszt, Lehrbuch des Deutschen Strafrechts, 22 Aufl., De Gruyter 1919.

von Liszt/E.Schmidt, Lehrbuch des Deutschen Strafrechts 1 Band, 26 Aufl., De Gruyter 1932.

中华人民共和国科学技术部「新一代人工智能伦理规范」（2020 年）

曾粤兴・高正旭「论人工智能技术的刑法归责路径」治理研究(2022 年第 3 期)

魏东「人工智能算法安全犯罪观及其规范刑法学展开」政法论丛(2020 年第 3 期)

（オンライン文献）

Brumfeld, Car assembly line robot kills worker in Germany, CNN (Jury 2nd ,2015)
https://edition.cnn.com/2015/07/02/europe/germany-volkswagen-robot-kills-worker/index.html
（最終アクセス 2022 年 11 月 24 日)

Car 2 Car Communication Consortium, Deployment of V2X communication based on IST-G5. https://www.car-2-car.org/fileadmin/press/pdf/CAR_2_CAR_Communication_Consortium_Statement_ITSG5.pdf
（最終アクセス 2022 年 11 月 24 日）

Car to Car Communication Consortium, Clear benefits for road safety and traffic efficiency.
https://www.car-2-car.org/about-c-its/
（最終アクセス 2022 年 11 月 24 日)

DARPA, AI Next Campaign, https://www.darpa.mil/work-with-us/AI-next-campAIgn
（最終アクセス 2022 年 11 月 24 日）

EUR-LEX, Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206
（最終アクセス 2022 年 11 月 28 日）

Mercedes-Benz Group, Mercedes-Benz is starting a Europe-wide cooperation project .https://group.mercedes-benz.com/innovation/case/connectivity/europe-wide-cooperation-car-to-x.html
（最終アクセス 2022 年 11 月 28 日）

Kühling, Roboter-Unfall bei VW in Baunatal: Angeklagter nach vier Jahren entlastet, HNA(18.05.2019)
https://www.hna.de/lokales/kreis-kassel/baunatal-ort312516/roboter-unfall-vw-baunatal-gutachten-entlastet-angeklagten-12294560.html
（最終アクセス 2022 年 11 月 28 日）

Kweitz, Ein Schweizer Bot im Darknet-Shopping-Wahn (und keiner weiß, wer haftbar ist), VICE (January 15th, 2013) https://motherboard.vice.com/de/article/78kyz4/random-darknet-shopper-590
（最終アクセス 2022 年 11 月 28 日）

Orgalim, European Regulation on Artificial Intelligence – Orgalim calls for legal clarity and workability, 21 April, 2021, https://orgalim.eu/news/european-regulation-artificial-intelligence-orgalim-calls-legal-clarity-and-workability
（最終アクセス 2022 年 11 月 28 日）

Wagner, Robot Liability, June 19, 2018, https://ssrn.com/abstract;3198764
（最終アクセス 2022 年 11 月 28 日）