

Title	AIから生ずる権利侵害と不法行為法
Sub Title	Les violations des droits découlant de l'IA et le droit de la responsabilité extracontractuelle
Author	白石, 友行(Shiraishi, Tomoyuki)
Publisher	慶應義塾大学大学院法務研究科
Publication year	2025
Jtitle	慶應法学 (Keio law journal). No.54 (2025. 2) ,p.[115]- 145
JaLC DOI	
Abstract	
Notes	片山直也教授・鹿野菜穂子教授・本郷亮教授退職記念号
Genre	Departmental Bulletin Paper
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AA1203413X-20250225-0115

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

AI から生ずる権利侵害と不法行為法¹⁾

白石 友行

はじめに

- I. 現時点における AI に関連した権利侵害と不法行為法
 1. 人の補助としての AI に関連した権利侵害と過失に基づく責任
 2. 人の補助としての AI に関連した権利侵害と過失の証明から切り離された責任
 - II. 近い (?) 未来における AI による権利侵害と不法行為法
 1. 人の代替としての AI による権利侵害と現時点の不法行為責任
 2. 人の代替としての AI による権利侵害と近い (?) 未来の不法行為責任
- おわりに

はじめに

AI は、現在さまざまな場面においてさまざまな形で用いられており、これからますます社会に浸透していくはずである。このことに伴って、AI は多く

1) 本稿は、2024 年 3 月 14 日に開催された Seminar on AI Technologies and Legal Theories “Innovation and Regulation” で行った “Artificial Intelligence and Civil Liability” と題する報告の一部、2024 年 10 月 7 日に開催された GenAI / SUM 2024（生成 AI サミット 2024）におけるワークショップ「責任ある AI の法学」で行った「生成 AI 社会の法的責任の所在とは」と題する報告の一部、および、2024 年 10 月 22 日に開催されたシンポジウム「AI による不正会計検知の現在と未来」で行った「不正会計検知 AI モデルの社会実装と法」と題する報告の一部を基礎にしている（ただし、最後の報告は本稿の脱稿後にされたものである。）。前二者について報告の機会を与えていただいた千葉大学の川瀬貴之教授およびリーテックス株式会社にあらためて感謝申し上げる。上記の経緯および紙幅の制約のため、本稿では文献の引用が最小限にとどめられている（文献の所在等については、注 2）で引用する拙稿を参照。）。また、本稿が扱う内容に関しては、近刊予定の別稿で詳細に検討することが予定されている。

の法的問題を生起させることが予想される。これらのうち現時点で取り組まれるべきもっとも重要な法的課題の一つは、AI との関連で生ずる権利侵害に起因して発生した損害をどのような形で填補するかという問題である²⁾。この問題を適切に規律することは、AI が社会に受け入れられるための前提条件になるからである。

AI との関連で生じた権利侵害または AI との関連で生ずる可能性がある権利侵害との関わりの中で求められるルールは、これに起因して発生した損害の填補についてのルールに限られない。AI との関連で生ずる可能性がある権利侵害を未然に防ぐためのルールも必要である。例えば、これまでさまざまな場で論じられてきた AI 倫理やガイドライン、および、EU の AI Act 等に代表される AI の規制に関わるルールは、主としてこの点に関わる。しかし、これらだけでは、AI との関連で生ずる権利侵害に関わるルールとしては十分でない。まず、AI 倫理やガイドラインは、現実には AI との関連で権利侵害が生じた場面についてはもとより、AI との関連で権利侵害が生じないようにする局面に関しても、一定の拘束力をもった対応をすることができない。また、AI 規制との関連でいえば、現在の日本には一般的な AI 規制法が存在しないだけでなく、仮に包括的な AI 規制法が存在する状態を仮定したとしても、その規制があらゆる AI のあらゆる側面に及ぶわけではないほか、規制に違反する形でさまざまな行為等がされることについては容易に想像することができるため、AI 規制法が存在するかどうかにかかわらず、AI との関連で権利侵害が生ずる事態

2) この問題に関わる筆者の論稿として、拙稿「AI の時代と不法行為法」千葉 37 巻 3 号（2022 年）1 頁以下、同「AI とフランスの民事責任法（1）（2・完）」千葉 38 巻 4 号（2024 年）232 頁以下、39 巻 1 号（2024 年）152 頁以下、同「AI と民事責任法——フランス法の検討を通じた日本法における課題の描写」『金山直樹先生古稀記念論文集』（日本評論社・2025 年）所収予定（本稿の脱稿の時点で同書のページ割が未確定であるため、この論文については、同論文の先頭からのページ数で引用する。）。AI を用いた金融取引に即した検討として、同「AI を用いた金融取引から生ずる民法上の諸問題に関する検討——SDGs の時代において自己決定と非差別を前提とした包摂を実現するために」千葉 38 巻 1 = 2 号（2023 年）328 頁以下、同「AI を用いた金融取引と民法理論——SDGs の時代における民法学の役割」武蔵野 20 号（2024 年）117 頁以下。

は避けられない。更に、不法行為法には権利侵害の予防機能があることからすれば、とりわけ AI 倫理やガイドラインが十分に固まっておらず、かつ、AI 規制法も存在しない現状のもとでは、AI との関連で生ずる権利侵害に起因して発生した損害の賠償を AI に関わる一定のアクターに課すことは、AI との関連で権利侵害が生ずる事態を減少させることに繋がる。これらの理由から、現在の状況を前提としても、また、AI 倫理やガイドラインが確立され、または、一般的な AI 規制法が用意された状況を想定しても、AI との関連で生ずる権利侵害に起因して発生した損害の填補のあり方を探求しておくことは必要不可欠である³⁾。

なお、このことは、さまざまな AI 倫理やガイドラインおよび EU の AI Act 等に代表される規制法が、AI との関連で生ずる権利侵害に起因して発生した損害の賠償という問題にまったく関係を持たないことを意味しない。これらは、民事責任の問題を直接的に扱うものではないが、これらで基礎とされているアプローチの仕方等には、民事責任の問題を考える際にも参考にされるべき部分が含まれている。例えば、さまざまな AI 倫理やガイドラインおよび規制法等で各アクターに要求されている各種の判断や行為等は、責任の基礎としての過失の中身を具体的に捉えたり、AI（または AI を組み込んだ製造物）の欠陥の有無を判断したりする際に一定の指針になる。また、EU の AI Act が基礎としているリスクベース・アプローチ、すなわち、AI に含まれるリスクを、許容することができないリスク、ハイリスク、限定的なリスク＝特定の透明性が要求されるリスク、最小限のリスクに分け、そのリスクの程度に応じて規制のあり方を変えるアプローチに関していえば、規制の局面において把握されるべきリスクと民事責任の局面において捉えられるべきリスクとでは質的に相違があり、前者におけるリスクに応じた AI の区別を後者にそのままスライドさせることは適切さを欠くのではないかという疑問との関連で一定の留保は必要であるものの⁴⁾、少なくともその思考枠組みそれ自体は、AI に含まれるリスクの程度

3) 大塚直「総括——科学技術の発展に伴う多様なリスクと不法行為法」NBL 1272 号（2024 年）56 頁。

に対応した責任ルールを要請するという意味で、民事上の法的責任のあり方を考えるに際して参考になる。

ところで、AI との関連で生ずる権利侵害に起因して発生した損害についてその填補を実現するためのルールに関しては、さまざまなものが想定される。これらのうち本稿が検討の対象にするのは、不法行為法を通じて AI に関わりを持つ人⁵⁾に責任を課すことにより当該責任主体に対して損害を賠償させるルールである。このことは、不法行為法以外の補償の仕組みを通じた損害の填補に関わるルールと AI 自体に責任を課すことによる損害の填補に関わるルールを検討の対象から除外することを意味する⁶⁾。

まず、不法行為法以外の補償の仕組みを通じた損害の填補に関わるルールを検討の対象から除外するのは、以下の理由に基づく。AI との関連で生ずる権利侵害に起因して発生した損害を填補する方法としては、大きく分けると、不法行為法によって特定の責任主体に損害を賠償させる方法と補償基金によって集団的に損害を補償する方法が想定される。従前の仕組みにおいて、補償基金による損害の填補は、主として、不法行為法によっては十分かつ迅速な救済を得ることができない被害者をそのままの状態に放置しておかないという政策的配慮により基礎づけられる。しかし、AI については、その判断や動作等を制御する可能性を持つ者やそこから利益を得る者が存在するため、AI との関連で生ずる権利侵害に起因して発生した損害に関しても、これらの者の責任を問うことによりその填補を図ることが可能である。従って、AI との関連で生ず

4) 小塚莊一郎「AI 製品に対応した EU の製造物責任ディレクティブ改正」情報法制研究 15 号（2024 年）48 頁で紹介されている Hacker の見解等。

5) 本稿は、議論の見通しをよくするために、AI に関わりを持つ人として、その開発者、提供者、利用者のみを想定する。AI に関わりを持つ人としては、これらの者以外にも、場合によって、例えば、データ提供者、学習者、クラウド提供者等が存在するほか、ある者が複数の役割を兼ねる場合や、複数の者が一つの役割を担う場合もある。

6) 拙稿・前掲注 2)「AI の時代と不法行為法」5 頁以下、同・前掲注 2)「AI とフランスの民事責任法（2・完）」138 頁以下、101 頁以下、同・前掲注 2)「AI と民事責任法」13 頁以下。

る権利侵害に起因して損害が発生する場面について、補償基金の創設を下支えしてきた政策的配慮はそのままの形では妥当しない。また、AI が浸透した社会ではその判断や動作等の正確性や安全性をどのように確保するかという点が重要な課題になる。しかし、補償基金の仕組みでは、損害の填補と帰責が分離されているため、AI に関わる人との関連でその判断や動作等の正確性や安全性を確保させることへのインセンティブが十分に働かない。これに対して、不法行為法においては、その予防機能を通じてこうしたインセンティブが適切な形で作用する可能性がある。これらの理由により、少なくとも現時点では、不法行為法の中で AI との関連で生ずる権利侵害に起因して発生した損害の賠償を実現する方法が探求されるべきである。

次に、AI 自体に責任を課すことによる損害の填補に関わるルールを検討の対象から除外するのは、以下の理由に基づく。AI との関連で生ずる権利侵害に起因して発生した損害の賠償を不法行為法によって実現する場合には、責任主体との関連で、AI に関わりを持つ人の責任を問題にする方法と AI 自体の責任を問う方法が想定される。AI 自体の責任という考え方は、2017 年頃に EU を中心に話題となったが、現在ではほとんど受け入れられていない。その理由は、民法の世界では人と物が明確に区別されているため、AI という物を人のように扱うべきではなく、また、この考え方によれば、法主体性の基礎づけ方によっては、一定の人に法主体性を認めることができなくなるおそれがあるなど、法理論的にみて、AI に法人格を付与することは適切でないこと、および、例えば、AI には賠償または付保の原資となる財産がなかったり、AI に関わりを持つ人が非責任化される結果、AI の判断や動作等の正確性や安全性を確保させることへのインセンティブが働かなくなったりするなど、AI に法人格を認めてその責任を問うことは実際上の意義に乏しく、場合によっては、有害ですらあることという二点に要約される。そもそも、現在では人間中心の AI 社会が目指されている以上、AI の責任ではなく、AI に関わりを持つ人の責任こそが問われなければならない。

このように、本稿は、AI との関連で生ずる権利侵害に起因して発生した損

害との関係で、AIに関わりを持つ人の不法行為責任がどのような形で問題になるかを検討する。AIとの関連で人の権利が侵害され、これによって人に損害が生じた場合、誰が、どのような理由に基づき、どのような要件のもとで、どの範囲で責任を負うか。これらが本稿で検討されるべき問いである。これらは、一見すると単純であるようにみえるものの、非常に複雑な問題である。そこで、これらの問題を検討するに際して強く意識しておくべきであると考えられるいくつかの基本的な理解の仕方または視点を、あらかじめまとめて提示しておくことが有益である。

第一に、AIを契機として権利侵害が生じた場合には、そのAIの開発、提供、利用の各段階でAIに関わりを持つ人の法的責任が論じられるべきであること、そして、AIでは学習やアップデート等の形でこれらのプロセスが一定の範囲で繰り返されるため、AIの開発者や提供者については当初の開発や提供の場面だけでその責任が問題になるわけではないこと、また、AIの利用者についてはその利用の前後における行為や判断等との関係でもその責任が問題になることである。このことは、AI以外の物や製品との関連でも基本的に当てはまるものの、AIとの関連ではその性質に起因してより顕著な形で現れる。

第二に、AIの自律性、学習性、不透明性等の特徴をしっかりと考慮して、この問題を論じなければならないことである。これらの特徴は、AI以外の物や製品にはみられない性格である。そのため、AI以外の物や製品との関連で生成されてきた不法行為責任のルールをAIとの関連でもそのまま妥当させることは適切か、そうでなければ、どのように既存のルールに修正や改良を施していけばよいかという視点を持ちつつ、AIに関わりを持つ人の責任の基礎、主体、要件、効果を考えることが要請される。

第三に、被害者の権利保障とイノベーションの促進をどのように調和させるかという視点をもって、AIに関わりを持つ人の責任のあり方を検討しなければならないことである。AIに関わりを持つ人に対して過剰な責任を課すと、AIの開発や利用が妨げられる一方で、これらの者に対して過少な責任しか課さないとすると、被害者の権利保障が十分に実現されず、ひいては、社会にお

ける AI への眼差しが悪化し、AI の利用にも悪影響が生ずる。従って、被害者の権利保障とイノベーションの促進のいずれにも配慮した適切な責任のあり方を探求することが求められる。

第四に、不法行為法の機能を考慮して、AI に関わりを持つ人の責任のあり方を検討しなければならないことである。不法行為法は、侵害された権利の回復および実現を目的とするだけでなく、権利侵害の予防という機能を持つ。そのため、AI による判断や動作等の正確性や安全性を確保し、AI を契機とした権利侵害の発生を未然に予防するためには、誰に対してどのような要件のもとでどの範囲の責任を負わせることが適切かについて考えておくことが必要となる。

本稿は、上記の四つの見方または視点を強く意識しながら、AI との関連で生ずる権利侵害に起因して発生した損害の賠償に関わる問題を検討し、責任原因に焦点を当てて、あるべき不法行為法上のルールについて一定の見通しを示す。その際、本稿は、「現時点における AI に関連した権利侵害と不法行為法」と「近い(?) 未来における AI による権利侵害と不法行為法」に分けて、AI との関連で生ずる権利侵害に起因して発生した損害の賠償に関わるルールを検討する。「近い」の後に括弧書きで「?」を付記したのは、そうした未来がいつかは来ると予想されるものの、それがいつになるかについて現時点では判断することができないからである。

本稿が検討の対象にする問題との関連で「現時点における AI に関連した権利侵害と不法行為法」と「近い(?) 未来における AI による権利侵害と不法行為法」を区別することには、二つの意味がある。一つは、AI の使われ方に関わる⁷⁾。AI は、現時点では、ほとんどの場面で、人の判断や行為等を補助するための道具として位置づけられている。しかし、AI の自律性や透明性などが向上する一方で、判断や行為等の主体およびそれらのプロセスよりも最終的な結果が重視されるようになれば、広く一般的に AI が人に代わって判断や動作等をすることが認められるようになるかもしれない。AI が人の判断や行為等の補助であるか、それとも、人の判断や行為等の代替であるかによって、AI

に関わりを持つ人の不法行為責任のあり方は大きく変わることが予想される。もう一つは、不法行為法のルールに関わる。AI が広く浸透した社会を見据えるのであれば、この問題を解決するために現時点における不法行為法のルールで十分か、未来のあるべき不法行為法はどのようなものかについても考えておかなければならない。

以下では、まず、現時点における AI の使われ方を前提とした場合に、現在の不法行為法によると、AI との関連で生ずる権利侵害に起因して発生した損害の賠償がどのように規律されるか、また、その規律にどのような限界があるかを明らかにする（Ⅰ）。次に、近い（？）未来における AI の使われ方を展望した場合に、AI による権利侵害に起因して発生した損害の賠償に関する規律として、現在の不法行為法にはどのような課題があり、この課題を解消するために、どのような形で不法行為法のルールを作ることが望まれるかを検討する（Ⅱ）。

Ⅰ．現時点における AI に関連した権利侵害と不法行為法

現時点において、AI は、ほとんどの場面で、人の判断や行為等を補助するための道具として位置づけられている。こうした AI との関連で権利侵害が生ずる場面としては、例えば、AI の判断や動作等が客観的にみて誤っていたものの、人がその誤りに気づかないまま当該 AI の判断や動作等を参考にしたリ

7) この区別について、米村滋人「AI 機器使用の不法行為における過失判断——医療・介護分野での責任判断を契機に」法時 94 巻 9 号（2022 年）48 頁以下、拙稿・前掲注 2）「AI の時代と不法行為法」9 頁以下等。また、人間の拡張と人間の代替、IA（知性の拡張）派と AI 派という視点を提示する、ジョン・マルコフ（瀧口範子訳）『人工知能は敵か味方か：パートナー、主人、奴隷——人間と機械の関係を決める転換点』（日経 BP 社・2016 年）、主体としての人工知能という見方と道具としての人工知能という見方を対置する、鈴木貴之「人工知能に関する 2 つの見方——主体としての人工知能と道具としての人工知能」同編著『人工知能とどうつきあうか——哲学から考える』（勁草書房・2023 年）3 頁以下等も参照。

利用したりして自己の判断や行為等をしたため、最終的にされた判断や行為等にも誤りがあり、その結果、他人に権利侵害が生ずる場面、AI の判断や動作等が客観的にみて誤っていたものの、人がその誤りに気づかないまま自己の判断や行為等を当該 AI の判断や動作等に委ねたため、最終的にされた判断や行為等にも誤りがあり、その結果、他人に権利侵害が生ずる場面、AI の判断や動作等は客観的にみて正しく、人が当該 AI の判断や動作等を参考にしたり利用したりしたものの、結局これとは異なる判断や行為等をした結果、最終的にされた判断や行為等に誤りがあり、その結果、他人に権利侵害が生ずる場面、および、AI の判断や動作等は客観的にみて正しかったが、人が当該 AI を利用せずに、これとは異なる判断や行為等をした結果、最終的にされた判断や行為等に誤りがあり、その結果、他人に権利侵害が生ずる場面等が想定される⁸⁾。現時点における不法行為法を前提とした場合、これらの各場面はどのように規律されるか。また、その規律には何らかの課題や限界があるか。更に、その課題や限界はどのようなものか。以下では、過失に基づく責任（1）と過失の証明から切り離された責任（2）に分けて、これらの問いを検討する。

1. 人の補助としての AI に関連した権利侵害と過失に基づく責任

人の判断や行為等の補助としての AI との関連で生ずる権利侵害に起因して発生した損害について、AI に関わりを持つ人が過失に基づく責任を負うかどうかを判断するためには、その前提として、AI の開発者、提供者、利用者のそれぞれについて、どのような義務が課されていたか、または、これらの者がとるべき適切な行動パターンはどのようなものであったかを明らかにしたうえで、その違反またはそこからの逸脱があったかどうかを検討しなければならない。I の冒頭で例示した各場面についていえば、人が何らかの判断や行為等をする際に AI の判断や動作等にそのまま依存すること、または、これに依存しないことは許されるか、AI の判断や動作等を参考にしたり利用したりして、

8) 金融取引との関連で想定される具体的な紛争事例について、拙稿・前掲注 2)「AI を用いた金融取引と民法理論」122 頁以下。

または、これを参考にしたり利用したりすることなく行われた判断や行為等はどうのように評価されるか、人の判断や行為等を補助するための AI の判断や動作等に何らかの誤りがあった場合において、人が当該 AI の開発、提供、利用、維持、制御等との関連で行った判断や行為等はどうのように評価されるか等の問いにつき、上記の枠組みに即した評価がされる。

まず、権利侵害をもたらす直接的な原因となった判断や行為等との関連では、AI がした判断や動作等それ自体の適否ではなく、人がした判断や動作等との関係で人に課されている義務または人に要求されている適切な行動パターンがどのようなものであったかが問題にされる。というのは、現時点における AI のほとんどは、基本的に人の判断や行為等を補助するための道具にすぎないため、この場合に法的な評価の対象になるのは、そうした AI の判断や動作等を参考にして、または、これらを参考にすることなく行った人の判断や行為等でなければならないからである。また、現時点における AI のほとんどは、その性能の制約等に起因して、あくまでも人の判断や行為等を助ける存在にすぎないことからすれば、たとえ客観的にみて正しい判断や動作等をする AI が存在したとしても、人が判断や行為等をするに際してそのような AI を利用しなかったことそれ自体を問題にして、または、これとは反対に、客観的にみて誤った判断や動作等をする AI があって、人がそのような AI を開発したり、提供したり、利用したりしたことそれ自体を問題にして、それらのことだけを理由に、開発者、提供者、利用者につきそれぞれの過失の有無を論ずることもできないと考えられる。以上によれば、AI の利用者については、AI の判断や動作等を補助として利用する形で、または、それらを補助として利用せずにとられた自己の判断や行為等が結果回避義務に違反していたかどうか、または、適切な行動パターンから逸脱していたかどうかにより、過失の有無が評価される。こうした過失の評価の仕方は、基本的には、AI と関わりなくされた人の判断や行為等に関する過失の評価の仕方と異ならない。その意味で、人の判断や行為等それ自体の法的な評価の仕方については、AI が補助手段として用いられているかどうかで大きな相違はない⁹⁾。

次に、権利侵害の直接的な原因となった人の判断や行為等が客観的にみて誤った AI の判断や動作等を参考にしたり利用したりしてされた場面では、当該人の判断や行為等との関係での過失のみならず、これらを補助した当該 AI の開発、提供、維持、制御等との関連で行ったその開発者、提供者および利用者の判断や行為等との関係での過失も問題になる。AI が人の判断や行為等を補助するための道具であるとすれば、その AI には、人による適切な判断や行為等を実現するのに適した補助手段であることが求められ、その開発者、提供者および利用者には、そのために必要となる一定の義務が課せられるからである。これによれば、AI の開発、提供、維持、制御等との関連で行ったその開発者、提供者および利用者の判断や行為等が、結果回避義務に違反していたかどうか、または、適切な行動パターンから逸脱していたかどうかによっても、AI の開発者、提供者および利用者の過失の有無は評価される。その際、学習性、自律性、不透明性等といった AI の特徴を考慮すると、開発者や提供者においては、いったん開発や提供をすればそれで義務が尽くされたと判断することはできず、当該 AI を適宜アップデートすることのみならず、これを追跡し調査したうえで、人の判断や行為等に負の影響を与える可能性との関連で必要があれば、適切な措置を講ずることが結果回避義務または適切な行動パターンとして措定され、その違反またはそこからの逸脱が認められると過失の存在が基礎づけられる。とはいえ、こうした過失の評価の仕方も、基本的な枠組みとしては、AI 以外の物や製品の製造者、提供者および利用者との関連で形成されてきたそれと大きく異ならない¹⁰⁾。AI の開発、提供、維持、制御等に関わる場面では、AI の特徴を反映してその利用の前後における行為規範の範囲が

-
- 9) 拙稿・前掲注 2)「AI の時代と不法行為法」12 頁以下、同・前掲注 2)「AI とフランスの民事責任法 (1)」222 頁以下、同・前掲注 2)「AI と民事責任法」3 頁以下。橋本佳幸「AI のリスクと無過失責任」NBL 1272 号 (2024 年) 39 頁以下も参照。
- 10) 拙稿・前掲注 2)「AI の時代と不法行為法」15 頁、同・前掲注 2)「AI とフランスの民事責任法 (1)」219 頁以下。医療の場面に関する記述であるが、角田美穂子ほか編『ロボットと生きる社会——法は AI とどう付き合う?』(弘文堂・2018 年) 428 頁以下〔米村滋人〕も参照。

拡大されているだけである。

ただし、こうした基本的な考え方には、いくつかの課題や限界を指摘することができる。まず、権利侵害をもたらす直接的な原因となった人の判断や行為等との関係では、これまでの裁判例等で蓄積されてきた判断枠組みを用いて過失の有無を評価することができるが、AIの開発、提供、維持、制御等との関連で行ったその開発者、提供者および利用者の判断や行為等との関係では、AIの用いられ方や特徴、そして、AIの判断や動作等を参考にしてとられた人の判断や行為等から生ずることが想定されるリスクの内容や大きさ等を十分に考慮に入れる必要があるため、AI以外の製品や物との関わりで形成されてきたその製造者や供給者等の過失に関する判断枠組みをそのままの形で用いることはできず、より具体的に、さまざまなAIとの関連で、誰がどの段階でどのような義務を負うかを明らかにしなければならない。この点の検討に際しては、さまざまな形で論じられてきたAI倫理やガイドラインの内容等も参考にされる。もっとも、こうした具体化の作業が必要であることは、過失の判断に共通する課題であり、AIの開発、提供、維持、制御等における過失の判断に固有の課題ではない。

次に、第一点との関連で各アクターにつき結果回避義務または適切な行動パターンの内容が明確にされたとしても、AIについては、開発、提供、利用に関わる者が数多く存在することも想定され、かつ、AIの判断や動作等の過程には不透明性が伴うため、誰がどの段階でどのような結果回避義務に違反し、または、適切な行動パターンをとらなかったかに関して、どのように証明すればよいかということも課題になる。これは、AIにおけるトレーサビリティの少なさにも関わる課題や限界である。

最後に、透明性と検証可能性のあるAIが開発され、AIへの信頼性が高まり、こうしたAIが相当程度にまで普及したときには、特にその利用者との関係でみると、二つの方向に過失の評価の仕方が変容する可能性がある。一方で、上記のような状況のもとでは、AIを利用すべきかどうか、また、どのAIを利用すべきかという次元における法的評価に影響が生ずる可能性もある。この場合、

特段の事情がない限り、補助手段として AI を利用すること、AI の判断や動作等を信頼して自らの判断や行為等を形成することがいわば適切な行動パターンの一つとして指定され、場合によっては、人が誤った判断や行為等をしたときに、客観的にみて正しい AI の判断や動作等に従わなかったことや客観的にみて正しい判断や動作等をする AI を利用しなかったことが過失の存在を強く推認させる要素として捉えられたり、客観的にみて誤った AI の判断や行為等に従ったことが過失の存在を否定する方向に作用する要素として把握されたりすることも想定される。ここでは、AI の性能の向上に伴い、AI への不信が低減され、AI への肯定的評価が高まるという理解のもと、そのことが、AI の利用者に課せられる行為規範の内容それ自体に変容を生じさせてはいないものの、過失の有無に関わる評価の次元で影響を及ぼしている¹¹⁾。他方で、上記のような状況が更に進んでいくと、AI が人の判断や行為等を補助する道具であるにとどまらず、それらの代替としての位置づけを与えられることも考えられる。この場合には、権利侵害の直接的な原因となった判断や行為等がそれ自体としては人との関わりを持たないため、それらとの関連で利用者の結果回避義務または適切な行動パターンを指定することもできない。その結果、AI に関わりを持つ人の過失に基づく責任が成立する範囲も狭まり、AI との関連で生ずる権利侵害に起因して発生した損害が過失に基づく責任によって填補される可能性も減少する¹²⁾。この点は、Ⅱ 1 で扱われる。

以上の検討によれば、現時点における過失に基づく責任に関するルールは、少なくとも人の判断や行為等を補助するための AI に関する限り、実体法の次元では、AI とは関わりなく生じた権利侵害に起因して発生した損害の賠償と

11) 穴戸常寿ほか編『AI と社会と法——パラダイムシフトは起きるか?』（有斐閣・2020 年）165 頁〔小塚莊一郎〕、167 頁以下〔森田果〕、拙稿・前掲注 2）「AI の時代と不法行為法」17 頁、同・前掲注 2）「AI を用いた金融取引と民法理論」143 頁、同・前掲注 2）「AI とフランスの民事責任法（1）」220 頁。

12) 橋本佳幸「AI と無過失責任——施設・機械の自動運転に伴う事故の危険責任・瑕疵責任による規律」法時 94 巻 9 号（2022 年）54 頁、拙稿・前掲注 2）「AI の時代と不法行為法」21 頁。

同じ程度には、AIに関連した権利侵害に起因して発生した損害の賠償を図ることができるルールとして評価される。とはいえ、AIの性能の向上とその社会への普及の程度等によっては、こうした評価に疑問が生ずる可能性もある。

2. 人の補助としてのAIに関連した権利侵害と過失の証明から切り離された責任

従前の議論では、AIを動物に準えてAIの占有者または管理者に対して動物の占有者または管理者に準じた責任を課す考え方、および、AIを被用者に準えて事業のためにAIを使用する者等に対して使用者責任に準じた責任を課す考え方が示されていた。しかし、これらの考え方は、AIをそれとはまったく異なる特徴を持つ動物に近づけて捉えたり、物であるAIを人である被用者として位置づけたりする点等において、適切さを欠く¹³⁾。現在の法において、人の判断や行為等を補助するためのAIに関連した権利侵害との関係でその適用が想定される過失の証明から切り離された責任は、製造物責任と運行供用者責任である。ただし、運行供用者責任に関しては、自動運転自動車に搭載されているAIの性能が人の判断や行為等の補助のレベルにとどまるかぎり、自動運転自動車の運行により他人の生命または身体が侵害された場面における同責任の適用との関係で深刻な課題や限界が新たに生ずるわけではない。運行供用者責任については、特に一定の条件下におけるレベル4の自動運転自動車またはレベル5の自動運転自動車の運行により他人の生命または身体が侵害された場面における同責任の適用との関連でいくつかの重大な課題や限界が存在するところ、これらのレベルの自動運転自動車は、人の補助としてのAIというよりも、人の代替としてのAIに関わるため、これらの課題や限界はⅡ1で扱われる。これに対して、製造物責任に関しては、AIが人の判断や行為等の補助のレベルにとどまったとしても、適用の可否、適用の範囲または領域、責任主

13) 拙稿・前掲注2)「AIの時代と不法行為法」24頁、同・前掲注2)「AIとフランスの民事責任法（1）」202頁以下、同・前掲注2)「AIとフランスの民事責任法（2・完）」118頁以下。

体、責任の範囲、免責等との関連で、いくつかの課題や限界がある。製造物責任については、とりわけ、以下の五つの課題や限界が指摘されなければならない。

第一に、製造物の意味に関わる課題や限界がある。製造物は製造または加工された動産とされ、かつ、動産は有体物に限定されているため、一般的な解釈によれば、情報やプログラム等は製造物に当たらない。この解釈によると、AI は、他の製造物の中に組み込まれているのであれば、製造物に該当しない。そのため、この解釈では、AI は他の製造物に組み込まれていないことが多いにもかかわらず、こうした AI の欠陥により生じた生命、身体、財産の侵害に伴う損害の賠償が製造物責任の対象外になってしまう。AI に関連した権利侵害との関係で製造物責任を広く適用するためには、解釈の変更や立法上の手当てを通じて、他の製造物に組み込まれているかどうかにかかわらずあらゆる AI が製造物に当たるとする理解の仕方が必要になる。

第二に、欠陥に関わる課題や限界がある。ここでは、ひとまず三点だけを指摘する。一つは、AI との関連で通常有すべき安全性をどのような基準で評価するか、特に人の能力を超える判断や動作等をする AI について通常有すべき安全性の基準をどこに求めるかという点である。一方で、合理的な人を基準として通常有すべき安全性を評価する方法は、AI に期待される役割等を考慮すると十分でなく、他方で、他の同種の AI を基準として通常有すべき安全性を評価する方法は、AI の性能の比較が難しいことを考慮するとやはり十分でないと考えられるからである¹⁴⁾。もう一つは、AI が通常有すべき安全性を欠いていたことをどの時点で判断するかという点である。AI については、その提供後においても開発者や提供者がその判断や動作等に影響力を持ち、また、一定のアップデート等も予定されているため、これまでの製造物と同じように引

14) 拙稿・前掲注 2)「AI の時代と不法行為法」20 頁の注 (57)、25 頁以下、同・前掲注 2)「AI とフランスの民事責任法 (1)」193 頁以下。刑事法に即した記述であるが、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」法時 91 巻 4 号 (2019 年) 22 頁以下も参照。

渡しの時点に限定する形で欠陥の有無を評価し、開発者や提供者による事後措置等を契機として何らかの不具合が生じたときにこれを欠陥とみなすという判断をするのは十分でないと考えられるからである¹⁵⁾。更に別の一つは、AIの欠陥により生命、身体、財産を侵害され、これにより損害を受けた者が当該AIの欠陥を証明することができるかという点である。AIには学習性や不透明性といった特徴があることに加えて、その技術的な難しさもあって、被害者がAIの欠陥を証明することには、ほかの製造物の欠陥を証明する場合に比べて、格段に困難が伴うからである¹⁶⁾。確かにデジタル証明等の技術が発展すればこの問題に一定の対応をすることは可能であるものの、何らかの法的な手当てをしておくことは不可欠である。これらの第二の諸点に関しては、EUの製造物責任指令の改正案でみられるように、AIの特徴を踏まえた欠陥の判断基準を定立すること、欠陥の証明に関する負担を軽減する措置を講ずることが必要である¹⁷⁾。

第三に、開発危険の抗弁による免責に関わる課題や限界がある。AIには予見することができない判断や動作等をする可能性があり、製造業者等であっても、引渡し時における科学技術に関する知見によってはAIに欠陥があることを認識することは困難であるため、AIとの関連では多くの場合またはほとんどの場合で開発危険の抗弁が認められかねないことが危惧されている。イノベーションの促進を重視する観点からこの抗弁をそのままの形で維持する選択

15) 拙稿・前掲注2)「AIの時代と不法行為法」26頁、同・前掲注2)「AIとフランスの民事責任法(1)」192頁以下。

16) 藤田友敬「自動運転をめぐる民事責任法制の将来像」同編著『自動運転と法』(有斐閣・2018年)283頁、東崎賢治＝松崎由晃＝羽鳥貴広「プラットフォームにおけるAI・IoTの活用とこれに伴う法的問題(下)」NBL1139号(2019年)76頁、拙稿・前掲注2)「AIの時代と不法行為法」27頁、同・前掲注2)「AIとフランスの民事責任法(1)」190頁以下。

17) 前者については、欠陥の具体的判断基準として、運用開始後に学習を継続する能力が製造物に与える影響、製造物が流通に置かれもしくは使用開始された時点、または、製造業者等がその後も製造物を管理している場合には製造物が製造業者等の管理下を離れた時点が挙げられている(6条)。後者については、証拠開示の仕組みが用意されているほか(8条)、一定の場合に欠陥の推定が認められている(9条)。

肢もある一方で、権利保障の観点を強調し、また、AI の特徴に鑑みればその学習性の代償である不確実なリスクに対する予防の要求がより高まるとして、製造物である AI については、開発危険の抗弁を認めないという立法的な態度決定やその適用範囲を限定するといった解釈をすることも考えられる¹⁸⁾。

第四に、保障される権利の中身や賠償の対象になる損害に関わる課題や限界がある。AI からは、自動運転自動車の誤作動や診療 AI の誤判断等の場面でみられるように、生命や身体への侵害に由来する損害のほか、第一の課題や限界が解消されることが前提になるものの、人の感情を把握する AI、プロファイリング等が利用された場面でみられるように、生命や身体への侵害に直接関わらない人格への侵害に由来する非財産的損害も生ずる。AI が広く浸透した社会においては、後者の損害も賠償の対象にしなければ、被害者の権利の保障は十分な形で実現されない。このこととの関連で、製造物責任が、生命、身体、財産の侵害により生じた損害だけを賠償の対象にしていることをどのように評価すべきか、その対象を人格の侵害により生じた損害にまで拡張すべきかが問われる。また、仮に後者の問いを肯定する態度決定がされたときには、第二の課題や限界との関連で、更なる検討課題が生ずる。この場合、通常有すべき安全性を欠いていることという、製造物の有体性および物理的な権利侵害を想定しているかのようにみえる欠陥の定義をそのままの形で維持することについて、疑問が生ずるからである。

第五に、責任主体に関わる課題や限界がある。まず、AI の開発や製造等に関わる者が多く存在することを踏まえると、どの範囲の者が責任主体としての製造業者等に当たるかという製造業者等の範囲に関係する課題に加えて、その開発や製造等に関わった者をどのように特定するかという証明の次元に関係する課題もある。次に、製造物責任の背後には、製造業者等はその製造過程を通じて製造物の欠陥に関わる危険を支配すべき立場にある以上そこから生ずる権利侵害に起因して発生した損害を賠償すべきであるという考え方があるとすれ

18) 拙稿・前掲注 2)「AI とフランスの民事責任法 (1)」189 頁以下、同・前掲注 2)「AI と民事責任法」7 頁。

ば、第二の課題や限界のうち二つ目の問題への対応の仕方にも規定されるが、AIはその提供者や利用者による学習の影響を受けることもあるため、AIに存在した欠陥のすべてを製造業者等による危険の支配に結びつけることはできるか、反対に、場合によってはAIに存在した欠陥をその提供者や利用者による危険の支配に結びつけるべき場面も存在するのではないか等といった疑問も生ずる。後者の疑問は、現在の法状況のもとでAIの提供者や利用者が基本的に過失に基づく責任のみを負うとされていることを踏まえるとより一層強化される可能性もある¹⁹⁾。

第一から第四までの課題や限界は、AIとの関連で生ずる権利侵害に起因して発生した損害が製造物責任によっては適切な形で填補されず、この責任による損害賠償のルールが被害者の権利保障のあり方として不十分と評価される可能性があること、とはいえ、製造物責任の基本的な枠組みを維持したまま、一定の解釈の変更または立法的な手当てをすれば、これらの課題や限界を解消することは不可能ではないことを意味している。これに対して、第五の課題や限界は、たとえ第一から第四までの課題や限界が解消されたとしても、製造物責任がAIの諸特徴を踏まえた責任のルールとして（ほとんどの場面では十分であるとしても）一定の場面では十分でないと判断される可能性があることを示唆している。

II. 近い(?) 未来における AI による権利侵害と不法行為法

近い(?) 未来において、AIは、多くの場面で、人の判断や行為等を補助するだけにとどまらず、これらを代替する存在として位置づけられる可能性がある。このとき、人はさまざまな場面でAIの判断や動作等に完全な形で依存することが許される。もとより、こうした状況が成立するかどうかは、判断や行為等それ自体との関連でどのような法的規律が用意されているか、具体的に

19) 拙稿・前掲注2)「AIとフランスの民事責任法(1)」196頁以下、同・前掲注2)「AIと民事責任法」9頁以下。

例えば、判断や行為等に人が介在することが法的に求められるかという点のほか、AIの信頼性と透明性がどの程度まで確保されているか、また、例えば、説明可能または解釈可能なAI等、AIの特徴に対応するための仕組みが用意されているかという点、義務や行為規範における重心が結果とプロセスのどちらに置かれているかという点、義務や行為規範の設定を通じて保護される権利はどのようなものかという点等に規定される。AIの信頼性と透明性が十分に確保されておらず、その特徴に対応するための仕組みが用意されていない場合は、予期せず、事後的に検証することができない形で、AIの判断や動作等による権利侵害が生ずる可能性が高まることから、これを回避するためにAIの判断や動作等に対して人の判断や行為等を介在させることが要求される。また、特定の結果よりも、人がどのように義務を履行し行為規範を遵守したかという点に評価の重心が置かれる場合には、AIに判断や動作等を任せるのではなく、人自身が判断や行為等をすることが前提になる。更に、義務や行為規範により保護されている権利が重要なものであるときには、その侵害が生じないようにするために、AIの判断や動作等とは別に人の判断や行為等を介在させることが求められる。これらの基準によれば、例えば、医業や法律事務をすることが医師や弁護士等に限定されているなど、法において特定の資格を持つ人による判断や行為等だけが許されている場面はもちろんのこと、現在のAIには不透明性や検証不可能性等の特性がありその信頼性も大きくないこと、AIからは生命や身体等のほか重要な人格の侵害も生ずること、現時点では結果のみならずプロセスも重視されていること等を考慮すると、少なくとも現在のAIの性能と法状況を前提にする限り、多くの場面では人がAIの判断や動作等に完全な形で依存することは許されないと考えられる。それでは、仮にこの状態に変化が生じ、人がAIの判断や動作等に完全な形で依存することが許されるという状況が成立した場合、人の判断や行為等の代替としてのAIによる権利侵害に起因して損害が発生した場合における規律として、現在の不法行為法にはどのような課題や限界があるか（1）。また、これらの課題や限界を解消するために、どのような形で不法行為法のルールを作ることが望まれるか（2）。IIでは、

これらの問いを検討する。

1. 人の代替としての AI による権利侵害と現時点の不法行為責任

近い（？）未来において現在の不法行為法がそのまま維持されていると仮定すると、人の判断や行為等の代替としての AI による権利侵害に起因して損害が発生した場合に、その AI に関わりを持つ人はどのような形でこれを賠償する責任を負うか。この点については、以下の見通しが得られる。

まず、人が AI の判断や動作等に完全な形で依存することが許される場合には、AI の判断や動作等が人の判断や行為等に代わり、AI による自律的な判断や動作等とは別に人による判断や行為等を想定することができないため、AI の利用者について、権利侵害の直接的な原因となった判断や行為等が利用者に課せられている義務または利用者に要求される適切な行動パターンに即したものであったかどうかを問うこと、従って、この部分の過失を理由として利用者に責任を課すことはできない²⁰⁾。その結果、AI の利用者に関しては、AI による判断や動作等の前後における過失を根拠とした責任の適用だけが問題になる。次に、AI の開発者や提供者については、AI が人の判断や行為等の補助にとどまる場合と同じく、その開発、提供、維持、制御等の各段階における過失を根拠とした責任の適用が問題になる。ただし、人の代替としての AI が用いられる場面においては、判断や行為等それ自体との関連で人の介在がないため、人の補助としての AI が利用される場面と比べて AI それ自体に起因する権利侵害のリスクが高くなることを考慮すれば、AI の利用の前後において各アクターに課される義務の程度または適切な行動パターンの基準が高くなることも予想される。

もっとも、人の代替としての AI が権利侵害を生じさせた場合のすべてにおいて、AI の開発、提供、維持、制御等の各段階における過失に基づく責任を各アクターに対して追及することはできない。というのは、実体法の次元でい

20) 「はじめに」で整理したように、権利侵害の直接的な原因となった判断や動作等をした AI 自体の責任を問題にすることもできない。

えば、人の代替としての AI による権利侵害に起因して損害が発生したからといって、あらゆる場面で、その開発者、提供者、利用者に上記の意味での過失が認められるわけではないからである。加えて、証明の次元でいえば、AI の判断や動作等の過程には不透明性があり、また、被害者側には技術的な知見の不足に由来する制約があるために、誰がどの段階においてどのような形で結果回避義務に違反し、または、適切な行動パターンをとらなかったかに関して、被害者側が証明することには困難が伴うからである。

更に、AI の開発者については、その解釈の変更や立法的な手当てがまったくされていないと仮定すると、他の製造物に組み込まれた人の代替としての AI が通常有すべき安全性を欠いていること、および、これにより他人の生命、身体、財産が侵害されたことが証明されたときには、製造物責任が課される。しかし、AI から生ずる権利侵害との関連で上記のように製造物責任を適用することに課題や限界があることは、I 2 で示したとおりである。しかも、そこで示した課題や限界のうち特に第二から第四までは、人の代替としての AI により権利侵害が生ずる場面では、より顕著な形で現れる。AI の自律性が高まれば高まるほど、通常有すべき安全性の基準が高くなることは予想されるものの、それと同時に欠陥の評価やその証明に関する困難も増大し、開発危険の抗弁に関わる問題はより顕在化するほか、自律的に判断や動作等をする AI が社会に浸透していくと、生命や身体に直接関わらない人格的な権利が侵害される場面も増えることが想定されるからである。

上記のような状態は、人の代替としての AI による権利侵害に起因して発生した損害の賠償に関わるルールのある方としては、十分なものとはいえない。AI が人から離れて自律的に判断や動作等を行うことが一般化した社会において、AI が人の判断や行為等の代替手段として機能するにすぎない社会における場合と比べて、AI に関わりを持つ人に対して過失を基礎とした責任を追及することができる場面が縮小するという状態、また、その開発者に対して製造物責任を問うことに関してより一層の困難が伴うという状態は、権利保障のある方として適切でない。確かに、こうした状態のもとでは、人の代替としての

AIに関わる各アクターが責任を問われるべき場面が明確な形で限定されるため、イノベーションの促進は実現されるのかもしれないが、それとバランスをとる形で達成されるべき被害者の権利保障に関しては十分な配慮がされていないというべきである。

ところで、人の判断や行為等の代替としてのAIのうち、特に一定の条件下におけるレベル4の自動運転自動車またはレベル5の自動運転自動車の運行により他人の生命または身体が侵害された場面では、運行供用者責任の適用が問題になる。ある者が自動車の運行供用者に当たるかどうかは、その者が当該自動車の運転を支配していたかどうかという観点からだけでなく、その者に運行支配と運行利益があったかどうかという観点から判断されるため、自動運転自動車にも運行供用者を観念することは一応可能である。従って、自動運転自動車の事故により他人の生命または身体が侵害され、これにより損害が生じた場合には、運行供用者責任の適用はある。そして、この責任によって被害者の権利保障は一定の範囲で実現される。もとより、特に完全な自動運転自動車の運行供用者の責任では、運転に関する注意義務と点検に関する注意義務のうち後者しか問題にならないこと、そのシステムのみで安全に運転することができる状態であったか、システムに異常が生じた場合にフェイル・セーフが組み込まれていたか等という点が問われることからすると、免責の判断が自動運転でない自動車の運行供用者の責任におけるそれとは異なることになるため、解釈論上の調整は必要になる²¹⁾。

しかし、運行供用者責任には、その適用領域に関わる限界がある。運行供用者責任は、AIを搭載した自動車以外の自動運転機械等により他人の生命または身体が侵害される場面、および、自動車の運行により他人の生命または身体以外の権利が侵害される場面には適用されず、AIを組み込んだ自動運転機械

21) 浦川道太郎「自動走行と民事責任」NBL 1099号（2017年）31頁以下、藤田友敬「自動運転と運行供用者の責任」同編著・前掲注16）134頁以下、後藤元「自動運転車と民事責任」弥永真生ほか編『ロボット・AIと法』（有斐閣・2018年）174頁以下、拙稿・前掲注2）「AIの時代と不法行為法」22頁以下。

等による生命や身体の侵害、および、自動運転自動車の運行により生ずる権利侵害のすべてを対象にしていない。また、運行供用者責任には、適用領域に関わる限界に加えて、自動運転自動車が普及した社会においても、この責任が現状のまま維持されるべきかという点に関わる課題がある。自動運転自動車による事故の原因の多くは、ヒューマンエラーではなく、システムの欠陥にあるため、この場合に、理論的には製造業者への求償の可能性が開かれているとしても、その製造業者ではなく運行供用者に対して一次的な責任を負わせることは十分に正当化されるか、むしろ、予防機能の観点からは、その製造業者に一次的な責任を課すことが適切ではないかという点が問われるからである²²⁾。

Ⅱ 1 における検討の結果は、以下のように要約される。人の判断や行為等を代替する AI により権利侵害が生じ、これによって損害が発生した場面においては、その開発者、提供者、利用者について過失に基づく責任を問うだけでは被害者の権利保障が十分に図られないおそれがある。AI 開発者の製造物責任を問うことにもいくつかの解釈論上または立法論上の課題がある。また、仮にこれらの課題が解消されたとしても、製造物責任には、AI の特徴を踏まえれば、ほとんどの場面で疑念は生じないものの、少なくとも一定の場面では、その提供者や利用者等による学習に影響を受けた AI の欠陥に結びつく危険を製造業者等としての開発者に帰すことはできるか、反対に、AI に存在した欠陥をその提供者や利用者による危険の支配に結びつけるべき場面も存在するのではないかという、その基礎づけに関わる課題が生ずる。このことは、運行供用者責任についても同じであり、そこでは、自動運転自動車の運行にほとんど影響力を持たない運行供用者に対して一次的な責任を課すことは適切か、反対に、自動運転自動車の運行から生ずる権利侵害についてはその製造業者による特別な危険の支配に結びつけるべきではないかという、その基礎づけに関わる課題が生まれる。加えて、日本の民法には、フランス法のように、物を原因とする一

22) 交通事故の責任主体として運転者と管理者を想定するフランスの 1985 年 7 月 5 日の法律においても、同様の問題が指摘されている。拙稿・前掲注 2) 「AI とフランスの民事責任法 (2・完)」152 頁以下。

般的な責任の法理が存在しないため²³⁾、解釈の修正や立法的な手当てが施されないことを前提とした製造物責任や運行供用者責任が適用されない領域では、AIに関わりを持つ人について過失の証明から切り離された責任を問うことができない。これらによれば、現在の不法行為法から導かれる帰結は、人の代替としてのAIから生ずる権利侵害に起因して損害が発生した場合のルールとしては、責任の基礎の観点からも、権利保障の観点からも、予防の観点からも十分ではない²⁴⁾。こうした認識を共有する立場からは、立法論として、AIによる権利侵害との関連で新しい不法行為責任のルールを作ることが要請される。

2. 人の代替としてのAIによる権利侵害と近い(?) 未来の不法行為責任

人の代替としてのAIによる権利侵害について、どのような理由に基づきそのAIに関わる人に対して責任が課せられるか、具体的に誰に対して責任が課せられるか、どのような要件のもとで責任が課せられるか、どの範囲の権利侵害または損害に関して責任が課せられるか、どれだけの期間この責任は問題になるか。これらは、立法論としてAIによる権利侵害に起因して発生した損害の賠償に関わるルールを考える際に解決されなければならない問いである。これらの問いのうち、責任の主体、要件、範囲、期間制限等をどのように理解するかという点は、責任の基礎をどのように捉えるかという点にすべてではないとしても大きく規定される。従って、AIによる権利侵害に起因して発生した損害の賠償に関わるルールを新たに構想する場合には、どのような理由に基づきAIに関わる人が過失の証明から切り離された責任を問われるかという点を考究することから始めなければならない。

不法行為法における帰責原理に関する従前の議論、および、AIから生じた

23) AIが他人に損害を生じさせた場合における物の所為に基づく責任の適用について、拙稿・前掲注2)「AIとフランスの民事責任法(1)」214頁以下。

24) この点は、日本私法学会第87回大会のシンポジウム「多様なリスクへの法的対応と民事責任——科学技術の発展に伴うリスクを中心に」で大塚直教授が示された認識と共通しているように思われる。

損害の賠償に関わる責任制度のあり方をめぐってフランスで展開されている議論を参考にすれば²⁵⁾、現在の不法行為法の基本的な体系との整合性を維持しながら、AI による権利侵害に起因して発生した損害をその AI に関わりを持つ人に対して転嫁するための根拠については、その者の過失を除くと、以下の三つの考え方が想定される²⁶⁾。なお、これらのほかにも、例えば、使用者責任に関する一般的な理解の仕方を参照して、その判断や動作等につき AI に不法行為責任が成立することを前提に、代位責任の枠組みに依拠してその AI の利用者等に責任を課す考え方も想定される。しかし、この考え方は、AI に法主体性を認めることを前提としており、この点において受け入れられないほか、責任の構造としてはともかく、実質的にみれば、物である AI の欠陥に関わる責任を問題にする考え方と大きな相違はないようにもみえることから²⁷⁾、本稿の検討対象から除外される。また、AI それ自体ではなく、AI を組み込んだ施設や機械等に関わる無過失責任を問題にするという態度決定もある²⁸⁾。しかし、人の代替としての AI がほかの有体的な物に組み込まれているかどうかにかかわらず広く自律的に判断や動作等をする社会を見据えるならば、施設や機械等に組み込まれた AI から生ずる権利侵害について過失の証明から切り離された責任を問題にするだけでは十分でないと考えられる²⁹⁾。

25) 拙稿・前掲注 2)「AI とフランスの民事責任法 (2・完)」114 頁以下、同・前掲注 2)「AI と民事責任法」3 頁以下。

26) 中原太郎「現代無過失責任論の一断面—— AI システムに起因する損害の填補をめぐる議論を素材として」法時 96 巻 8 号 (2024 年) 45 頁以下は、表題の議論について、危険責任論、古典的な危険責任論によらない嚴格責任、代位責任の活用、補償基金の活用に分けて整理する。これらのうち二番目の考え方は、不法行為責任の体系の再編も要請されるため (46 頁)、本文で述べた基本方針に従って検討の対象から除外される。また、三番目および四番目の各考え方も、「はじめに」で述べたことに従い検討の対象にはされていない。

27) 中原・前掲注 26) 49 頁。

28) 橋本・前掲注 12) 54 頁以下、同・前掲注 9) 33 頁以下。

29) 大塚・前掲注 3) 54 頁は、AI を組み込んだ施設や機械等に関わる無過失責任を問題にする考え方について、施設の稼働者にそのコントロール可能性とは関わりなく責任を負わせること、極度に観念化した支配概念の活用により AI のブラックボックス性が追認される可能性があることを問題点として指摘する。

第一に、AI に存在する欠陥を防止するのに適した立場にある者は、そこから生ずる権利侵害に起因して発生した損害について責任を負わなければならないという考え方が想定される。この考え方は、製造物責任の基本的な枠組みを前提にしながら、I 2 で示した第五の課題や限界を意識して、その責任主体を製造業者から AI に存在する欠陥を防止するのに適した立場にある者一般へと拡張する。この考え方によれば、ほとんどの場合には、AI の開発者が責任主体になる。というのは、AI の開発者は、プログラムの設定等を通じて AI の欠陥を防止するのに適した立場にいるからである。ただし、AI の提供者や利用者等も、AI に与える学習の内容やその程度等によっては AI に欠陥を生じさせることがあるため、場合によっては、責任主体になる可能性を持つ。また、欠陥および保護の対象になる被侵害権利の捉え方次第では、この責任で対象にされる AI の範囲に結果として一定の限定が付される可能性はあるものの、この考え方それ自体においては、あらゆる AI の欠陥に起因して発生した損害がこの責任の対象に含まれる。

この考え方には、後述する第三の考え方にみられるのと同じ利点がある。しかし、なお検討されるべき課題もある。この考え方は、既存の製造物責任との連続性において AI による権利侵害に起因して発生した損害の賠償に関わる責任を構想する。そのために、この考え方においても、AI による権利侵害に起因して発生した損害の賠償との関連で製造物責任を適用する場合に生ずる課題や限界への応答が必要になる。確かに、この考え方では、I 2 で示した製造物責任に関わる課題や限界のうち、第五のそれは解消されており、また、AI に関する特別な不法行為の規律を用意するという前提に立つ以上は、第一のそれも解消されている、つまり、ほかの有体的な物に組み込まれていない AI もその対象になるという態度決定がされている状態を前提にすることができる。これに対して、第二から第四までの課題や限界については、立法的な応答がされなければならない。加えて、特に第二の課題や限界との関連では、高度の自律性や学習性を備えた AI が、予期しない形で、また、事後的に検証することができない形で人の権利を侵害する可能性があることを想起したとき、そして、

欠陥の評価と過失の評価の近接性も併せて考慮したときに³⁰⁾、欠陥を責任の基礎としつつ、その判断基準をどのように設定するか、また、その証明の困難にどのように対応するかという問題設定で十分か、これらの問いを超えて、より本質的に、欠陥という要件を排除すること、従って、この責任の基礎それ自体を放棄することが適切ではないかという点が問われる。

第二に、AI の判断や動作等を決定しそれらから利益を得ている者は、その AI による権利侵害に起因して発生した損害について責任を負わなければならないという考え方が想定される。この考え方の背後には、AI の判断や動作等は人の利益のために行われるという認識がある。この考え方によれば、責任主体として、AI の判断や動作等からもっとも直接的に利益を得ている者、具体的にいえば、その使用者が措定される。場合によっては、その所有者が責任主体になる可能性もある。また、人が AI の判断や動作等から利益を享受することについてはあらゆる AI との関連で想定することが可能であるため、この責任の対象には、その自律性の程度やリスクの大きさ等を問わずすべての AI による権利侵害に起因して発生した損害が含まれる。

しかし、この考え方に対しては、いくつかの重大な問題を指摘することができる。ここでは、三点のみを取りあげる。まず、AI による権利侵害との関連でのみ、しかし、権利侵害の原因となったあらゆる AI との関連において、その判断や動作等から利益を受けているという理由だけで、その自律性の程度やリスクの大きさを問うことなく、その使用者、場合によっては、その所有者に対して、責任の範囲に限定のない無過失責任を課すことは正当化されるかという疑問がある。これは、利益の帰属という観点からの責任の正当化がこの場面に妥当しないのではないかと、および、こうした正当化が看過しえない責任範囲の肥大化をもたらすのではないかとという疑問である。

次に、AI の使用者、場合によっては、その所有者に対して責任を課すだけでは、AI による判断や動作等の正確性や安全性を確保すること、そして、AI

30) 欠陥と過失の関係、製造物責任の法的性質に関わる議論について、窪田充見編『新注釈民法（15）』（有斐閣・2017 年）629 頁以下〔米村滋人〕等。

による権利侵害を適切な形で予防することができないのではないかという疑問がある。これは、そのプログラムの設定等を通じてAIの判断や動作等に直接的な影響を与えることができるのはその開発者等であるが、この考え方ではAIの開発者等の責任を問うことができないため、不法行為法の機能を通じた権利侵害の予防を実現することができないのではないかという疑問である³¹⁾。

最後に、この考え方のもとでは、AIの判断や動作等からもっとも直接的に利益を受けている者が一義的に確定されないこともあるため、誰がAIによる権利侵害に起因して発生した損害について責任を負うのかも明確にならないことがあるのではないかという疑問がある。責任主体があらかじめ明確にされていなければ、AIの判断や動作等によって権利を侵害された者は誰を相手方として損害賠償を請求すればよいかが分からず、このことは、被害者の権利保障の妨げになる。また、この問題は、潜在的な責任主体として措定されるAIの利用者にも関係する。AIの利用者において、自己がAIによる権利侵害について責任を問われるかどうかをあらかじめ明確に知ることができないことがあるとすれば、この利用者は、(AIによる権利侵害に起因する責任が保険の対象になるかどうかはひとまず措くとして)その責任につき保険契約を締結すればよいかを判断することもできなくなる。

第三に、AIまたはその判断や動作等に関わる特別のリスクを支配し制御する可能性を有している者は、そのAIによる権利侵害に起因して発生した損害について責任を負わなければならないという考え方が想定される³²⁾。この考え方の背後には、一定のリスクを含むAIは人によって制御されなければならないという認識がある。この考え方によれば、AIの開発者が一次的な責任主体になる。というのは、AIの開発者は、そのプログラムの設定等を通じてAIの判断や動作等に直接的な影響を及ぼすことができるからである。ただし、AIの提供者や利用者等がAIに与えた学習の内容やその程度によっては、これら

31) より一般的な文脈での指摘であるが、大塚・前掲注3) 53頁以下。

32) 拙稿・前掲注2)「AIとフランスの民事責任法(2・完)」110頁以下、同・前掲注2)「AIと民事責任法」3頁以下。

の者も AI のリスクを支配しているとみることができるため、場合によっては、AI の提供者や利用者等も責任主体になる。また、この考え方においては、リスクの支配という観点が重要な意味を持つため、この責任の対象には、すべての AI ではなく、危険責任を正当化することができるだけの特別なリスクを持つ AI による権利侵害に起因して発生した損害だけが含まれる。この考え方は、EU の AI Act 等で示されているリスクベース・アプローチに親和性を持つ。

この考え方には、第二の考え方との対比でみると、いくつかの利点がある。まず、この考え方は、現在の不法行為法における無過失責任の規律と一定の連続性を持つ。現在の不法行為法に存在する無過失責任の多くは危険責任の発想により基礎づけられる。また、危険責任は、特別の危険がある場合に正当化され、特別の危険は、危険の高度性と制御不可能性の程度により判断される。一定の AI には両者がともに存在し、また、危険の高度性が抽象的にはともかく具体的に明らかでない AI に関してもこれまで危険責任が妥当してきた場面上に強い危険の制御不可能性が認められるものがあるため、少なくとも一定の範囲の AI による権利侵害に起因して発生した損害の賠償に関わる責任については、危険責任の発想に基づく正当化が十分に可能であると考えられる³³⁾。また、危険責任により正当化される無過失責任は、これまで、主に生命や身体が侵害される場面を対象にしてきたため、危険責任の発想に基づく責任を AI により生命や身体以外の権利が侵害される場面に及ぼすことは難しいのではないかという点についていえば³⁴⁾、人格的な価値の重要性および AI による人格権侵害の著しさやその頻度を考慮したとき³⁵⁾、一定の人格権という留保を付すかどうかは別として、AI との関連では人格権侵害の場面を含む形で危険責

33) 中原・前掲注 26) 47 頁以下は、広い範囲の AI システムを対象にするのであれば、危険の高度性と制御不可能性をセットで嚴格責任を正当化する古典的ケースとは異なり、後者を起点とする新たな危険責任類型の定立が求められるのではないかとしたうえで、それがいかにして可能かは現状では明らかでないとする。

34) 中原・前掲注 26) 46 頁。

35) 拙稿・前掲注 2) 「AI の時代と不法行為法」33 頁以下、同・前掲注 2) 「AI を用いた金融取引から生ずる民法上の諸問題に関する検討」328 頁以下。

任の発想に基づく責任を構想することも選択肢になる。以上の検討によれば、AIによる権利侵害に起因して損害が発生した場合でこの考え方をを用いることには、既存の無過失責任法理との連続性を一定の範囲で確保するという意義が認められる。加えて、この考え方によれば、一定の特別のリスクを含むAIによる権利侵害に起因して発生した損害だけが責任の対象になるため、AIの開発者、場合によっては、その提供者や利用者に対して過剰な責任が課せられることもない。その結果、一定の限度という留保は付さなければならないとしても、第二の考え方と比べれば、過剰な責任のためにAIの開発や利用が妨げられるおそれも小さくなる。

次に、AIまたはその判断や動作等に関わる特別のリスクを支配し制御する可能性を有している者に対して責任を課すことは、AIによる判断や動作等の正確性や安全性を確保し、AIによる権利侵害を適切な形で予防することに繋がる。というのは、AIの開発者、場合によって、その提供者や使用者は、相対的にみてAIの判断や動作等を制御するのに適した立場にあり、これらの者に対してAIによる権利侵害に起因して発生した損害を負担させ、リスクを内部化させることを通じて、適切な予防を図ることが可能になるからである³⁶⁾。

最後に、この考え方のもとでは、AIの判断や動作等によって権利を侵害された者は、第二の考え方における場合よりも容易に損害賠償請求の相手方を特定することができる。確かに、AIの開発者、場合によっては、その利用者がたくさん存在する場合には、多くの潜在的な責任主体が存在するため、権利侵害を受けた者が誰に対して損害賠償を請求すればよいか分からなくなる事態も生じうる。しかし、この問題は、制度の仕組みを工夫すれば、一定の範囲で解消される。例えば、責任に順位付けをしたり、ひとまず特定のアクターに責任を集中させたりする方法がそれである。こうした工夫を施すことは、被害者の権利保障をより実効的なものにするという観点からも望ましい。

Ⅱ 2 の検討によれば、少なくとも近い（？）未来におけるAIの使われ方を

36) 大塚・前掲注3) 53頁以下の分析を参照。

展望した場合、人の代替としての AI による権利侵害に起因して発生した損害の賠償に関する責任の基礎としては、第二の考え方は適切さを欠く一方で、第三の考え方が相対的にみて望ましいと考えられる。

おわりに

人の補助としての AI を前提にする場合、AI との関連で生ずる権利侵害に起因して発生した損害については、人の判断や行為等それ自体に関わる義務の違反および AI の判断や動作等の前後における人の義務の違反を理由とする責任により、AI が関わらない場面と同じ程度の賠償を実現することができる（Ⅰ 1）。しかし、この場面で製造物責任を適用することにはいくつかの解釈論上または立法論上の課題や限界があるため、これらが解消されないのであれば、現在の不法行為法で十分な権利保障が実現されるかどうかには若干の疑問が生ずる（Ⅰ 2）。これに対して、人の代替としての AI を前提にする場合には、AI による権利侵害に起因して発生した損害について、現在の不法行為法では適切な形で賠償することができない（Ⅱ 1）。そのため、危険の支配という観点からその構造と証明に関するルールを基礎づけ、かつ、AI の開発者等に加えてその提供者や使用者をも責任主体として想定した AI 責任法により、AI による権利侵害に起因して発生した損害を賠償することが望まれる（Ⅱ 2）。この責任と運行供用者責任や製造物責任との関係を含め、その具体的なルールの中身を明確に提示すること、また、この責任を補充する形で補償基金や国家による填補等が必要になるかどうかを検討し、具体的な制度の全体像を明確にすることが次の課題である。