

[論文]

# 人工知能の罪

佐藤 英明

〈目次〉

- はじめに
- 1 感情
  - 1-1 応報感情
  - 1-2 ルールと感情
  - 1-3 釈明としての謝罪
- 2 責任
  - 2-1 多義性
  - 2-2 因果
  - 2-3 釈明と責任
- 3 意志
  - 3-1 動物裁判
  - 3-2 行為の原因としての意志
  - 3-3 自由における因果性
- おわりに

## はじめに

自動運転技術に関して、自動運転車が死傷事故を起こした場合の責任の帰属が検討課題となっている。自動車事故については操作する運転者の刑事責任が問われるが、完全自動運転の場合、運転操作をおこなう人間は存在しない。人間に刑事責任を帰属できないとすれば、自動運転車両そのものあるいはそれに搭載された AI (artificial intelligence) の刑事責任を問えるか否かが問題となる。

違法行為に対する非難という刑罰の応報的機能という観点から、違法行為を惹起させたプログラムの削除が自動運転車両に対する社会の応報感情に即した対応となるという見解に対し、AIにとって刑罰が一定の負担と感じられるものでない以上、AIに人格を肯定し刑罰を科す意義は想定しえないし、「『AIを処罰できるか』という設問それ自体が、一般人からは突飛な発想と受け止められるであろう」という指摘もある。他方、応報の要素や道義的非難から離れ、刑法の犯罪予防という観点だけから考えるならば、事故を起こした AI に刑事責任を認め、AIを破壊したりプログラムを消去したりすることは危険の除去につながるとみることでもできるが、非難要素を含まないそうした予防目的のみに基づく処分（保安処分）を刑罰と呼ぶことはできないとされる<sup>(1)</sup>。また、伝統的責任概念からも AI の刑事責任は否定される。

だが、人の死傷という重大な結果が生じているにもかかわらず、誰も刑事責任を負わないといった事態は、社会的に受容されず、特に被害者や遺族にとって受け入れがたいものとなるだろう。AIの刑事責任という問題は、刑罰そのものの意味の再検討を促す契機ともなる。本稿では、AIの罪について考えるために、感情、釈明、責任、意志といった観点から処罰について考察する。

# 1 感情

## 1-1 応報感情

自動車事故の刑事責任については、運転者にとって事故の回避が可能であったことが前提とされる。しかし、運転者のいない自動運転車両の場合、乗員が事故回避措置をとることは不可能であり、乗員の刑事責任を問うことはできない。自動車メーカーや自動走行システムの設計者の刑事責任を問うことは不可能ではないが、現実には刑事責任が肯定される可能性は低いと考えられている<sup>(2)</sup>。人間が運転している場合であれば過失運転致死傷罪が成立するような事案でも、不運な天災に遭ったかのように取り扱うしかなくなる。これを社会的に容認できるであろうか<sup>(3)</sup>。

運転者の刑事責任を問えない場合に問題となる帰責の間隙を社会がどのように受けとめるべきかは、極めて困難な問題であるといわれる。「交通事故が発生し、被害者の死傷という結果が生じたにもかかわらず、何人も刑事責任を問われないということが社会的に受容されるのかという問題」<sup>(4)</sup>は、そもそも刑罰のあり方を社会がどう受け入れるかという問題でもある。運転自動化システムが高度化すれば、人間が運転に関与する領域は減少してゆく。それとともに自動車運転過失致死傷罪の成立範囲は縮小する。しかし、人の死傷という重大な結果が生じているにもかかわらず、誰も刑事責任を負わないということは、被害者や遺族にとって受けいれがたいものとなる予想される<sup>(5)</sup>。それゆえ、自動運転車の事故の犠牲者を天災に遭ったかのように扱うことは、社会的に受容されない可能性もある。

この問題を考える際の一つのポイントは、被害者やその家族の「感情」であろう。交通事故に対する刑事責任については、被害者遺族などの意見を反映し、いくつかの法改正が行われてきた。刑法38条1項は「罪を犯す意思がない行為は罰しない」と規定し、故意犯を処罰することを原則としている。

しかし、同項には但書があり「法律に特別の規定がある場合は、この限りではない」とし、特に危険な行為や保護法益が大きい犯罪については、過失犯でも処罰されるとしている。他人の生命にかかわる危険な行為である自動車の運転には高い注意義務が課せられており、過失により人を死傷させた場合でも処罰の対象とされる。かつては、業務上過失致死傷罪（刑法221条）により処罰されていたが、飲酒運転によるひき逃げなど悪質な運転による事故が多発したことにより、被害者の遺族や国民から、飲酒運転や無謀運転、無免許運転など運転者の悪質な行為による事故に対し業務上過失致死傷罪の刑では「軽すぎる」という意見が出された。これを反映し、2001年に「故意犯」処罰の規定として「危険運転致死傷罪」が新設され、最高20年の有期懲役が科されることになった<sup>(7)</sup>。

また、2007年には「自動車運転過失致死傷罪」が新設され、「自動車の運転上必要な注意を怠り、よって人を死傷させた者」には、通常の業務上過失致死傷罪より重い刑罰が科されることになった。さらに、2013年には「被害者遺族をはじめ交通事故事犯の厳罰化を求める国民の声を受けて」、準危険運転致死傷罪が新設されるとともに、運転の悪質性、危険性の実態に応じた処罰を可能にするため、これらの犯罪行為については刑法から切り離され、新たな法体系である「自動車の運転により人を死傷させる行為等の処罰に関する法律（自動車運転処罰法）」が運用されることになった。

自動車事故に関するこうした厳罰化の動きは、被害者遺族や国民の声に促されたものであり、交通事故に関する処罰を社会がどう受け入れるかという問題に被害者等の「感情」が大きくかかわっていることを反映している。自動運転車の事故に関する刑事責任について検討する際にも、それは大きな問題になると考えられる。

他方、近代の刑事司法制度は応報感情による制裁の弊害を克服することを目指してきたといわれる。だが、克服の対象であった応報感情を重視することは、それに逆行するものとも受け取られる。そのため「刑事司法の原則を一般市民の感情レベルにおいて浸透させる」ための「感情教育」の必要性が

<sup>(9)</sup> 説かれたりもする。応報感情は理性が支配すべき法の世界においては、排除されるべき非合理的なものとも考えられている。

## 1-2 ルールと感情

戸田正直は、野生環境を背景として遺伝的に基本枠が設定された行動選択・行動遂行のための心的ソフトウェアを「アージ・システム」と呼び、「感情」もそのうちのひとつとみなしている。たとえば「恐れ」や「怒り」のような感情と結びつく逃走、威嚇、攻撃といった「状況対処行動」は動物にもみられるが、こうした行動を引き起こす仕組みを感情の起源とみるならば、感情は人類の進化とともに複雑化してきた状況対処行動のシステムと考えることができる。このような観点からみれば、感情の働きには「合理性」を見出すことができる。「人間という新種が数百年前に地球上に出現した時は、感情は人間にとっても合理的な存在であったと考えるのが自然である」<sup>(10)</sup>。

しかし、「文明化」された現代社会において、感情が非合理的な側面をもつことは否定できない。環境状況に応じて適切な状況対処行動を選択させることで個体が「生き延びる」ことを助けるのが、感情の本来の機能であった。感情システムは、長大な時間をかけて「生き延びる」ために有利に働くように進化をとげてきた。ところが、人間は「文明化」によって、初期人類時代の野生環境から環境状況を大きく変化させた。この変化は、感情システムの進化という時間スケールからみると、極めて短い時間のうちに生じたものである。そのため、「文明環境」においては、感情システムは必ずしも「生き延びる」ために有利に働くことにはならず、合理的とはいえ行動に結びついてしまうこともあると考えられる。

戸田は「怒り」という感情について考えるうえで、動物の「縄張り(territory)」防衛行動がヒントになるという。多くの動物は、生き延びるための資源占有空間を確保するために「縄張り」を作り、同種の他個体の侵入から防衛する。縄張りは、「臭いづけ」などの方法により表示されるが、その中に同種他個体が侵入することがある。占有者は縄張り防衛のために、威

嚇や攻撃などの行動をとる。こうした行動の背後にあるものを、「怒り」という感情の祖型（原怒り）とみることができる。

「縄張り」は、種が遺伝的に共有する行動の「ルール」である。こうしたルールは無益な闘争を避けるという意味で有用である。臭いづけマーカ―を認知せずたまたま侵入してしまった他個体に対する威嚇行動は、「ルール違反」への「警告」として機能する。この警告によっても立ち去らない侵入者に対して攻撃行動が起こる。縄張りが侵害されると、占有者は侵害者に対する「怒り」の感情を起動する。それによって、その感情特有の行動の準備状態としての身体的変化が生じる。「怒り」の場合、攻撃行動の準備状態として、表情が険しくなる、身体の筋肉の緊張により震えるといった身体変化が起こり、それが侵害者に対する威嚇となり「警告」となる。それによって侵入者が立ち去れば、闘争は避けられる。しかし、それが「意図的なルール破り」であった場合には、侵害者を追い出すために「罰」を与える必要がある。そして縄張りがルールとして意味を持つためには、他個体の縄張りを尊重する行動傾向が種全体で共有されている必要がある。それゆえルールを守り守らせようとする「感情」はとくに重要なものとなる。

人間は言語能力を持つため、「警告」は言語によっておこなわれることになる。それでも動物の縄張り防衛に見られる「原怒り」と人間の「怒り」の感情の間には連続性があると戸田はみる。他個体の縄張りに侵入した動物の退去は、知覚的に確認できる。しかし、人間が他者の権限を侵害した場合には、その侵害からの「退去」を目に見えるかたちで示すことができないこともある。そこで、自分は権限を侵害したが「これは誤りでしたので退去致します。ですから私を攻撃しないでください。」という信号を言語的に表明する。これが「謝罪」である。侵害の程度が軽い場合には、被侵害者の「怒り」はそこで収まる。

しかし、侵害に実質的損害が伴うときには、謝罪だけでは「怒り」は収まらない。その場合、強い「怒り」を取束させるには、受けた損害の「補償」とともに、侵害者に対する「攻撃」としての「侵害者加罰」が求められるこ

となる。野生環境においても、侵害者にルール違反を繰り返させないためには、違反の程度に応じた罰が与えられる。そして「怒り」の程度に比例する量的加罰は、「現代においても法執行機関が与える刑罰の基本理念になっている」という。

他方で、違反者に「罪悪感 (guilt)」のような感情が生まれれば、「罰を甘受して被害を補償しようとする行動傾向」に結びつく。それゆえ、被害者の「怒り」と加害者の「罪悪感」の働きがうまく噛み合えば、「社会秩序維持機能は正常に作動する」こととなる。<sup>(11)</sup> わかりやすくするために動物の行動も現代の人間社会をモデルに説明されているが、「怒り」という感情について「ルール」「謝罪」「罰」との関係が戸田は以上のように論じている。

### 1-3 釈明としての謝罪

人間には「不祥事が起これば、責任者を探し出し罰したいとする願望」があり、その背後には「すべてのことには原因があるはずだという信念」があるといわれる。「良いことをすれば報われ悪いことをすれば罰せられる世界にわれわれは生きている」という信念は、心理学では「公正世界信念 (belief in a just world)」<sup>(12)</sup> と呼ばれる。これによって、悪いことが起こったからには、その「原因」があるはずであり、悪いことを引き起こした人は「責任」を負って罰せられなければならないと信じられることになる。それは、責任を負うべきものを探し出して罰したいという「応報願望」<sup>(13)</sup> をもたらす。人は応報願望から負事象 (悪いこと) に関連したものを追求しようとする。ルール違反や他者への危害などの負事象との関連が問われた人間は、さまざまな言語的説明をおこなう。そうした説明は「釈明 (account)」と呼ばれる。「謝罪」は、言語的説明としての「釈明」の一種と考えられる。

釈明には、謝罪のほか、弁解、正当化、否認があるとされる。「否認」は、「わたしはやっていない」というように、被害をもたらしした行為を自分がおこなったことを否定することである。「正当化」は、自分の行為が被害をもたらしことは認めるが、「わたしは規則に従ってそうしただけだ」というよ

うに、その行為の不当性を否定し、非難の対象となることを否定することである。「弁解」は、自分の行為が被害をもたらしたこともそれが非難の対象となるものであることも認めるが、その非難（罰）が自分に向けられることは否定することである。「上司の指示によっておこなったことだから、上司が責任を負うべきである」というように、自分がおこなったことにより負事象が生じたことやそれが非難の対象となるものであることは認めるが、「責任」を負うのは自分でないと主張するのである。それに対し、「謝罪」は、加害行為への関与とその不当性、責任のすべてを肯定するものである。それゆえ謝罪するためには、加害行為への関与およびその不当性を認め、自分が責任を負うことを「表明」する必要がある<sup>(14)</sup>。しかし、それだけでは謝罪として十分とはいえない。言語表現だけをみた場合、謝罪の要素は表1のようになる。

表1 謝罪の要素<sup>(15)</sup>

謝罪要素	意 味	表現例
負事象の認知	自分の行為によって被害などの負事象が発生したことを認める。	「ご迷惑をおかけしました」 「こんなことになっていたとは知りませんでした」
責任受容	負事象の発生に対して自分に責任があることを認める。	「それは自分の責任です」 「そうなったのは自分のせいです」
悔悛表明	自分が悪かったと認識し、反省していることを表明する。	「申し訳ないことをしました」 「反省しています」
被害者へのいたわり	被害者の苦しみに理解を示し、これを和らげようと努める。	「たいへん気の毒に思っています」 「辛い思いをされたことと思います」
更生の誓い	危害や違反行為を繰り返さないと誓う。	「二度としません」 「これからは決してしません」
赦しを請う	被害者（あるいは関係者）に赦しを求める。	「どうか赦してください」 「勘弁してください」

先述のように、「謝罪」は加害行為への関与とその不当性を認め、自分が責任を負うべきことを肯定することである。それゆえ「責任受容」は、謝罪と他の釈明との違いを示す決定的な要素である。さらに謝罪において必須の要素となるのが「悔悛表明」である。責任を認めても、「申し訳なく思って



いる」といった悔悛の表明がないと、その謝罪は本心からのものではないと感じられ、「謝罪の真正さ」が疑われることになるからである。それゆえ、謝罪の中核的要素は責任受容と悔悛表明であるとされる。

責任受容と悔悛が謝罪において重要な理由は、被害者の関心に求められる。被害者にとって大きな関心事は損害が償われるか否かである。加害者が責任受容を表明すれば、損害賠償を期待できる。それにより、賠償に対する不安や懸念が解消される。また、被害者が「自分が悪かったのでは」といった自責や後悔などの気持ちを持っていた場合、加害者が「非難されるべき」であることを表明することにより、そうした自己批判的な感情から解放される。加害者が行為の改善を誓うことで、被害者の不安が鎮められるという効果もある。さらに、自ら非を認め謝罪することは、自らを惨めな立場に置くことであり、体面を傷つけられる屈辱的な行為と感じられることが多いため、加害者に対する社会的罰という意味もある。そのため、被害者の処罰願望を満たし、怒りや敵意を和らげることにつながる<sup>(16)</sup>。

このような点で、加害者による謝罪は、被害者にとって有益である。しかし、その謝罪が被害を償い被害者の苦痛を癒したいという気持ちや罪悪感からなされたものではないと疑われれば、被害者の不安や懸念、自責や後悔、怒りや敵意を解消することには結びつかない。それゆえ、「謝罪の真正さ」の根拠となる悔悛表明も、謝罪の中核的要素の一つとなる。加害者が謝罪した場合、被害者がそこに悔悛の気持ちを感じるほど赦そうという気持ちが強まることは、社会心理学の実験によっても確認されている<sup>(17)</sup>。

AIに「釈明」という言語的説明をおこなわせることは可能だろう。対話型AIは人間との「会話」のなかで謝罪の言葉を表明することがある。そこに「悔悛の気持ち」を感じとることはできるだろうか。

## 2 責任

### 2-1 多義性

刑法における規定は、日常生活における規範や正義に関する人々の理解や確信と密接に結びついている。刑法に定められた刑罰は身近なものではないとしても、罰は日常生活のうちにあり、われわれはどのようなものが罰せられるべきかを判断している。罰のイメージはわれわれの日常生活のうちに深く根ざしている。その意味で「刑法は、我々にとって遠くに、同時に、近くに存在する」<sup>(18)</sup>。

日常生活における罪は何を意味しているのだろうか。「罪」という語は、宗教、道徳、法律などの規範に「そむく行い（不正行為）」を意味するとともに、そうした不正行為に対する「罰」や「責任」の意味でも用いられる。「罪を犯す」という場合は「不正行為」を、「罪に服する」「罪を減じる」では「罰」を、「罪を他人にかぶせる」の場合は「責任」の意である<sup>(19)</sup>。規範にそむくことを行えば、責任を問われ、罰せられる。「罪」という語は、こうした一連のプロセス全体を意味しており、そのうちのどの部分に重きが置かれるかにより、意味の違いが生じると考えることもできるだろう。そして、参照される規範が刑罰法規の場合、規範にそむく行為は「犯罪」とされる。

「責任」という語は、中国語の「責（しなければならない）+任（任務）」が語源とされる<sup>(20)</sup>。しかし、現在この日本語はたんに「責務」という意味で用いられているわけではない。法学における「責任」という語の多義的な用法の原因として、瀧川裕英は「継受元のドイツにおいて複数の語で言い表されていた概念を、日本では『責任』の一語に担わせたこと」をあげている。「日本語の『責任』に対応するドイツ語としては Schuld ; Haftung ; Zurechnung ; Last ; Verantwortung ; Verantwortlichkeit などを挙げることができる。同様に英語においても responsibility ; liability ; guilt ;

accountability ; answerability ; burden ; imputationなどを挙げるができる<sup>(21)</sup>。以下では、瀧川による「責任」概念の分析を参照し、その多義性を整理しておく。

responsibility という単語ひとつをとっても、その意味は多様である。ハートは、responsibility という語の多義性を示すために次のような例をあげている。この例のなかの「責任のある (responsible)」という語は、それぞれ異なる意味で用いられている<sup>(22)</sup>。

船長として、Xはその乗客と乗組員の安全に対して責任があった。しかし、前回の航海で彼は毎晩酔いつぶれており、定員いっぱいの乗客を載せた船の遭難に責任があった。彼は正気ではないという噂が流れたが、医師は彼が自らの行為に対して責任があると判断した。航海の間中、彼はまったく無責任に行動していたし、これまでのさまざまな経歴からも彼が責任ある人物ではないということが明らかだった。船長は、遭難の責任は異常な冬の嵐にあると繰り返し主張したが、刑事訴訟において、彼は自らの過失行為に対して刑事責任があるとされた。それとは別に、民事責任訴訟において、生命と財産の損失に対して法的責任を負わされた。彼は今でも生きており、多くの女性と子どもの死に対して道徳的に責任がある。

責任が問われる状況には、「過去に関する責任状況」と「未来に関する責任状況」がある。前者は、すでに何らかの負事象（不利益・侵害等）がもたらされた状況である。この状況において責任が問われるには、なんらかの規範に対する違反が前提となる。前者が、あくまでも「事後責任」であるのに対して、後者は「事前責任」であり、責任を負うべき者を事前に指示することにより予防措置を期待するものである。前者が事後的に負担（不利益）を負う負担責任であるのに対し、後者は、ある立場の者に対しその立場（役割）を占めることで発生する何らかの責務を課す責務責任である。ハートの例では「船長は乗客と乗組員の安全に対し責任がある」というのが、責務責任である。「無責任に行動していた」「責任ある人物」という用法もこれと関連する。

前者の「過去に関する責任状況」においては、ある存在者（答責者）がある規範（責任規範）に違反したとき、その事態（責任対象）について、他の存在者（問責者）に対して、負担（責任負担）を負うことになる。責任が問われる際に参照される規範の種類により、「刑事責任」「民事責任」のような法的責任と道徳的責任を区別することができる。そして、この区別に対応して、刑罰、損害賠償、社会的非難などの責任負担が区別される。上記の例における「刑事責任」「民事訴訟における法的責任」「道徳的責任」がこれにあたる。「釈明」は、責任を問われた答責者の問責者に対する言語的説明であるが、その一種である「謝罪」は、答責者が責任を負うことを認め問責者の赦しを請うものであるとともに、答責者にとっては責任負担という意味をもつこともある。

## 2-2 因果

「過去に関する責任状況」において、すでにもたらされた結果に対する「原因」を指摘するために「責任」という語が用いられることもある。これが「因果責任」である。上記の例では「船の遭難の責任は異常な冬の嵐にあった」「船長は船の遭難に責任があった」がこれに該当する。嵐が遭難の原因という場合、日本語ではふつう「嵐のせい（所為）」と言うことはあっても「嵐に責任がある」という言い方はしないが、英語の responsibility は、人間に対しても自然現象に対しても用いられる。「責任」という語が「非難対象」に対して使用されるのに対し、responsibility は「せい（所為）」と同様にたんなる「原因」についても用いられる。

因果責任があるものに負担責任が認められるためには必要な要件がある。責任能力と呼ばれるような精神的条件もそのひとつである。上記の例では、「自らの行為に対して責任がある」という医師の判断が、それに関係している。それゆえ、「因果責任があるもの＝負担責任を負うもの」とはならない。

自動運転車のような AI を搭載した機械が事故を起こしたとき、「因果責任」はその機械にある。AI に刑事罰を科することができるかという問いは、

因果責任のある機械に刑法規範に基づく負担責任を負わせることが可能かという問いと捉えることができる。この問いについて最初に考えなければならないのは、行為と出来事の違いである。

瀧川は、「ある出来事の生成原因である」という意味での責任を「生成責任」と呼び、「一応の非難対象者である」ことを意味する「有責責任」と区別する。有責責任の有無と密接に関係するのが、生成責任が帰属されるのが「出来事」か「行為」かという違いである。「自動車事故の責任はブレーキの故障にある」という場合には、ブレーキの故障という「出来事」に生成責任が認められるが、「自動車事故の責任は運転者の前方不注意にある」という場合には、運転者の「行為」に生成責任がある。前者は「出来事生成責任」、後者は「行為生成責任」と呼ばれる。出来事生成責任の探求は、「原因」の探求であり、原理上、無限に遡ることが可能である。事故原因はブレーキ故障、その原因は部品の劣化、その原因は劣悪な使用環境……というように限りがない。それに対して、行為生成責任の追求は無限に遡ることはなく、その帰着点は「行為者」である。行為者に生成責任（行為生成責任）が認められたとき、その行為者に有責責任が帰属させられる可能性が生じる。出来事生成責任から行為生成責任が区別されるのは、それが有責責任の帰属の前提として要請されるからである。

それゆえ、AI搭載機械の振る舞いを「行為」と見なすことができるとすれば、AIにも有責責任があるといえる可能性がある。しかし、有責責任の帰属にはさまざまなレベルがあり、帰属条件によって有責責任の有無は異なる。生成責任さえあれば有責責任もあるとされる場合、予見可能性が必要とされる場合、意図が必要とされる場合、さらに責任能力が必要とされる場合など、その条件によって有責責任の帰属は異なる。

瀧川は「関与責任」を「過去の出来事に対する何らかの作用・生成・連関・関与を意味する概念」と定義する。関与責任を認めるということは、出来事の原因をある存在者に帰属させるということである。関与責任は「負わされたり引き受けられたり果たされたりする」ものではなく、あるかないか

を認定されるものであるため「ある責任」とも呼ばれる。関与責任は、生成責任と有責任に区別され、さらに生成責任は出来事生成責任と行為生成責任に区別される。そして、ある存在者がある事態に対して行為生成責任があると認められた場合に、その存在者に有責任が帰属される可能性が生じることになる。

それに対して「規範違反の結果として発生する負担あるいは不利益」を意味するのが「負担責任」である。負担責任については「責任を負う」という表現が用いられる。関与責任が「ある責任」であるのに対し、負担責任は「負う責任」である。法的には刑罰や損害賠償が、道徳的・社会的には非難や配慮・辞任・解任などが、負担責任の内容となる。

関与責任や負担責任が、すでに何らかの事態が生じた「過去に関する責任状況」において問題となるのに対し、「未来に関する責任状況」において果たされるべき責務は「責務責任」と呼ばれる。「人がある立場・地位・役割を占めることで発生する何らかの責務」であり、「親は子供の養育に責任がある」といった場合が、これにあたる。運転者の「安全運転義務」も責務責任とみなすことができる。負担責任が、あくまでも「事後責任」であるのに対して、責務責任は「事前責任」であり、何かが起こったときに責任を負うべき者をあらかじめ指示し、そのようなことが起こらないように予防措置をとることを期待するものである。

したがって、「責任」という概念は表2のように整理することができる。

表2 責任の概念<sup>(23)</sup>

概 念		意 味	
関与責任 (ある責任)	生成責任	出来事生成責任	出来事の生成原因であること
		行為生成責任	行為の生成原因であること
	有責任		行為者が一応の非難対象であること
負担責任 (負う責任)		負担 (あるいは不利益) ……事後責任	
責務責任		責務 ……事前責任	

### 2-3 釈明と責任

前述の「否認」「正当化」「弁解」「謝罪」という釈明の分類についても、以上のような「責任」概念の区別と関連づけて説明することができる。「わたしはやっていない」といった否認は、「自分には行為生成責任がない」と釈明していることになる。なんらかの害が生じているが、その生成原因は自分の行為ではないということである。「規則に従っただけ」といった正当化は、行為生成責任は認めるが「自分には有責任がない」という釈明である。法的には「違法性阻事由」に対応する。「上司の責任だ」といった弁解は、関与責任（行為生成責任+有責任）は認めるが「自分の負担責任は否定される」という釈明である。責任能力のない子供がもたらした被害に対する損害賠償責任を親が負うことになるといった場合、子どもの関与責任は肯定されるが負担責任は否定される。法的には「責任阻却事由（免責事由）」に対応する。それに対し「迷惑をかけて申し訳ない」といった謝罪は、関与責任（行為生成責任+有責任）と負担責任を認め、赦しを請う釈明ということになる。

有責任の帰属条件の一つとされるのが「制御可能性（controllability）」である。それゆえ、これは「弁解」という釈明と関わり、制御可能性が高いほど、責任は大きいと判断される。たとえば、弁解の効果に関する実証的研究では、「電車に乗り遅れたから」といった制御可能性（回避可能性）の高い弁解よりも、「母親を病院に連れて行かなければならなかったから」といった制御可能性の低い弁解の方が好意的に受け取られ、弁解の効果が高いことが確認されている<sup>(24)</sup>。

制御可能性が最も高いと考えられるのは「意図的な危害」であり、他方、最も低いとされるのは「不可抗力による危害」である。その間の段階は連続的に変化する評価次元となるが、「意図的危害」と「不可抗力」の間に「怠慢」「不注意」という段階を設定することもできるし、「意図的」「非意図的」の二段階に分けることもできる<sup>(25)</sup>。責任との関係において自由意志が問題とな

るのは、このように有責任の帰属条件となるためでもある。

### 3 意志

#### 3-1 動物裁判

先に述べたように AI を搭載した機械の振る舞いを「行為」と見なすことができるとすれば、AI の有責任を肯定できる可能性がある。では行為と出来事はいかにして区別されるのか。刑法において「犯罪」とみなされるのは「行為」だけである。そして犯罪は「人」の行為であるとされ、自然現象や動物に刑罰を科すことには意味がないとされる<sup>(26)</sup>。

しかし中世ヨーロッパでは、動物裁判がおこなわれていたことが知られている。現代ならば事故や自然災害として処理されるような出来事が、動物の犯罪行為として正規の裁判所で、人間とまったく同じ訴訟手続きと厳正な審理を経て判決が下され、刑が執行された<sup>(27)</sup>。現代人の目には、このような裁判は意味のない不合理なものと映るだろう。大屋雄裕は、われわれが動物裁判を無意味なものとする理由を「行為指導性」という観点から説明している<sup>(28)</sup>。人間は、法に反する行為が処罰の対象となることを予測して、そうした行為への関与を事前に避けることができる。そのため、法制度には、特定の行為を人々に行わせたり行わないようにさせたりする「行為指導性」がある。しかし、法を理解できず処罰を予測できない動物に対してそのような事前規制は不可能である。それゆえ野生のイノシシが住宅内に立ち入っても住居侵入罪は成立しない。イノシシの「住居侵入」は「行為」とは認められないのである。

法による事前規制が意味をもたないという理由から動物や自然現象を「犯罪」から排除するとすれば、同じ理由によって一部の人間も「犯罪者」として扱うことはできなくなる。結果を予期し自らの意志によりそれを回避できるような能力を持たない人間の行動を規制することはできない。責任能力と



呼ばれるような精神的条件が処罰の前提条件とされている理由もまた、「行為指導性」という観点から理解することができる。しかし、われわれは責任能力を欠いた人間の行動については「出来事」ではなく「行為」と考えているのではないか。動植物や無生物に行為者性を認めない理由が、行為指導性がないという点だけだとすれば、人間の行動であるというだけでは「行為」とは認められないことになる。だが、われわれは行為指導性を出来事と行為の区別の基準としているだろうか。人間と人間以外の存在というより根源的な区別を前提として、行為は出来事と峻別されているのではないか。

現代人から見て不合理と思われる動物裁判は、理性と対置されるような処罰願望や応報感情の暴走、不合理な擬人化などのためにおこなわれたのであろうか。『動物の刑事訴追と死刑 (The criminal prosecution and capital punishment of animals)』は、1906年に出版されたエドワード・ペyson・エヴァンズの著作である。この著作は、1988年にニコラス・ハンフリーの緒言をつけて復刻されたが、その抄訳が『殺人罪で死刑になった豚』(1995年)である。記者の遠藤徹は「解題」において、この本の「発掘者」であるハンフリーが動物に対する処罰がおこなわれた三つの理由をとりあげていることに言及している<sup>(29)</sup>。第一の理由は、社会的な危険をなくすため。第二の理由は、動物に戒めを与え再発を防ぐため。第三の理由は、人間がそれを見て罪を犯さないようにするためというものである。しかし、危険をもたらす動物は殺して排除すれば裁判にかける必要はないし、動物に対する戒めの効果は期待できない。人間が罪を犯さないようにするには人間の裁判で十分である。それゆえ、これらを動物裁判がおこなわれた理由とみることはできない。現代人が処罰の根拠と考えるような理由では説明できないのである。さらにハンフリーは、「人を殺した剣」や「人を押し潰した銅像」などの無生物までもが裁判にかけられたという事実をあげる。そして、次のように結論づけている。「理不尽な出来事は、世界の秩序が失われることとして人々を激しい不安に陥れたのである。そこで人々は、法廷という人間世界に秩序を回復する場において非合理的な出来事を裁くことで、世界に秩序と合理性を回

復しようとしたのではないか」。

池上俊一も、動物裁判はけっして非合理的な考え方に基づいていたわけではなかったという。<sup>(30)</sup> 動物裁判がおこなわれるようになった12・13世紀は、「神盟裁判や裁判決闘のような非合理的な証明システムにかわって、合理的な証拠裁判主義があらわれた」時代である。それまでの「呪術的世界観を前提とする神聖な秩序に生じた傷の反射的・固定的修復にかわって、犯人の意図や動機の反省・倫理的基準にもとづく非難と責任追及をとまなう『刑罰』が、誕生したのである」。「善悪は『人格』とその『意図』があってはじめて生じる倫理的関係であり、その責任も、その『人格』がおうのだ」という考え方が、その当時の「哲学・神学の新しい傾向」であった。

こうした考え方の背景として、12世紀に自然が考察の対象となり、それをいかにして合理的な用語で説明していくかが課題となりはじめたことが関係していると池上はみる。それ以前、「森」に代表されるような自然は、不可思議な驚異にみちた「異空間」であり、法や秩序にみすてられた者、自ら世を捨てた者のすみかであり、普通の人間には容易に近づけない場であった。しかし、人間が自然の征服をすすめ、人間と自然の関係を捉えなおそうとするようになると、自然は合理化され「脱魔化」されるようになる。それまで、負事象の源泉は、森のような神秘的空間にひそむ魔力に求められてきた。だが、自然が合理化され、魔力を失うようになると、それは、負事象をもたらした当事者の「人格」に内属するものとされる。魔性は、森という自然のうちにあるのではなく、個々の人格のうちにあると考えられるようになったのである。人格の内面としての良心や意図が倫理や法の課題とされるようになってきた時代だったからこそ、魔女や動物にも「人格」を認め、悪行の「原因」が悪しき「意図」に求められたのではないかと池上は指摘している。

### 3-2 行為の原因としての意志

ケルゼンは、ギリシア語で「原因」を意味する *aitia* という語の原義が

「罪 (guilt)」であることを指摘している。<sup>(31)</sup> *aitia* は「責務」「責任」を意味する語であったが、そこから転じて「原因」を意味するようになった。この語は、もともと「ある事態がなぜ生じたのか」という問いに対する説明となるものを意味する言葉であった。なんらかの負事象が生じたとき、「原因」が追求されるのは、「それがなぜ生じたのか」という問いに対する答えになるからである。負事象の生成への関与を疑われたものはその事態が生じた理由の説明(釈明)を求められる。先に述べたように、「罪」という語は、規範にそむくことを行い、責任を問われ、非難され罰せられるという一連のプロセス全体を意味している。*aitia* の追求は、非難され罰せられるべき対象の特定である。非難すべき人間が特定されれば、*aitia* の追求はそこで終わる。「行為者は行為の最終原因と見なされ、行為者を超えて因果関係を遡らな<sup>(32)</sup>い」。

それに対し、出来事の「原因」の連鎖には限りがない。自然現象や動物の行動の原因は、際限なく遡ることができる。出来事の因果連鎖のような「自然因果性」についてしか語ることができないとすれば、非難すべき対象として *aitia* を追求することは不可能となる。*aitia* を追求するためには、非難すべき行為は出来事とは区別されなければならない。

しかし、人間の行為であっても、それを出来事として捉えることはできる。黒田亘は「出来事の因果連鎖」と「行為の因果連鎖」を対比し、両者は「同じ一つの世界についての二つの語りかた」<sup>(33)</sup> であるとした。ある人の「右手が上がる」という出来事は、その人が「右手を上げる」という行為でもある。生理現象としてみれば、その原因は腕の筋肉の収縮、中枢からの神経伝達、それをもたらした感覚刺激、さらにその原因と、「どこまでも、いわば不定無限に遡ることができる」。しかし、同じ動作を行為としてみれば、その因果連鎖は行為者自身から発現し、「別の行為者には結びつかない」。行為の因果連鎖を原因の方向にむかってたどれば、主体となる存在によって区切られるが、出来事の原因の連鎖は無際限に<sup>(34)</sup>続いている。

アリストテレスは意図的行為(本意からの行為)とはその始まり(原因)が

行為者自身のうちにある行為であると規定した。<sup>(35)</sup>近代になると、デカルト、ホッブズ、ロックなどによって、行為は意志が原因となって身体動作が引き起こされることにより成立するという「古典的意志理論」が広まる。<sup>(36)</sup>「行為と行為でないものの区別の基準を意志の作用という内面的な過程の有無に求める」という考え方である。「先行する意志の働きによって惹き起こされた行動、それが行為であり、そうでない行動は、外見はどうであろうと行為ではない」と広く考えられるようになった。しかし、ウイトゲンシュタインは「私が自分の腕を上げるという事実から、私の腕が上がるという事実を引き去るとき、後に残るのは何か」<sup>(37)</sup>と問うた。これは意志の働きという原因を想定する古典的意志理論に対する問いかけである。

黒田によれば「行為と行為でないものの区別あるいは限界は、自然 (physis) に基づくのではなく、いわば人間社会の約束ごと (nomos) に属している。『行為』という概念は、人間が生きてゆく仕組みの一部として、長い人間の歴史を通じて沈殿、定着したものである」<sup>(38)</sup>。「行為」という言葉が用いられるのは、「非難」の対象となる場合、「責任」の問題が論じられる場面、規則や規範との関係においてとくに「違反」が問題とされる場合である。前述のように、「罪」とは、規範にそむくことを行い、責任を問われ、非難され罰せられるプロセスのことである。つまり「行為」について語られるのは、「罪」に関連する場面である。「ある人間的な現象を行為と見なすこと」は、それを「『規則』『価値』『責任』等々の概念の適用対象であるもの」と考えることである。われわれは「意志の働き」など確認しなくても、行為と行為ならざるものとを区別している。他者の心の中に「意志」を見出さなくても他者の「行為」について語ることができるし、行為者自身が自分の「意志」を意識していなくても「行為」をなすことはできる。それゆえ「『行為』の定義的基準とされる意志過程なるものは、あらかじめ常識の了解によって行為ならぬ現象から区別されている人間の営みの背後に、ことさら仮定された内的過程であり、たいていは架空の存在なのである」。

### 3-3 自由における因果性

中島義道は、カントの「自由における因果性」における「超越論的自由」とは、「日常生活において普通に行為するときに誰でも遭遇する事柄」であるという<sup>(39)</sup>。それは、ある身体運動を行為として認めるなら、その行為記述と同一の意志記述を認めなければならないという根源的な了解のことである。「殴る」という行為には「殴ろう」という意志が、「食べる」という行為には「食べよう」という意志が論理的に伴っているという因果性が「自由における因果性」である。これは「空虚な論理的要請」ではなく、自然因果性に基づく記述を行為の記述として認めず、行為記述から排除する機能をもつ。「行為」とは、「たんに現実に『なされた』行為」としてではなく、「『なされるべき』行為」という観点から把握されるものであり、なんらかの規範的原理との関連においてのみ意味をもつものである。「行為」がなんらかの規範との関係においてしか語ることでできないものであるとすれば、自然因果性に基づく記述によって行為を語ることはできない。

そして、規範を前提として行為を語るということは、責任について語ることでもある。自由な意志は、自然因果性における原因として行為に先行しているのではない、責任を負うべき結果が生じた後で責任が問われるとき、責任追及の終止点とされるのが自由な意志であり、その意味で自由な意志は「責任の原因」である。行為者に責任を負わせる際に、行為記述（行為の説明）のなかで事後的に行為の原因としての意志を「構成」するのである。それゆえ釈明という言語的説明においても、「意図」や「意志」に言及されることになる。動物裁判において悪行の原因が「悪しき意図」に求められたことも、動物の行為の原因の「構成」と理解することができるだろう。

中島は、意志過程のような「心の状態」よりも「責任追及のほうがより根源的である」という。たとえば自動車事故が起こると、「わき見運転」とか「注意義務違反」などの「原因」が探り当てられる。事故原因は「運転手の心の状態」であるかのように記述される。しかし、それは彼の「心の状態そ

のもの」を表しているのではない。「道路の脇を歩いていた小学生の集団に突っ込んで数人を跳ねてしまった運転手の男が制限速度をオーバーしておらず、飲酒運転もしていなかった、携帯電話に夢中であつたわけでもなく、無免許運転をしていたわけでもない。すると——しかたがないので——彼には『注意義務違反』というレッテルが貼られる」。しかし、「注意義務違反」は「なまの心の状態（自然的事実）」ではなく、「社会的要求に基づいた事実」である。その意味では、一種の「フィクション」であるが、それを「あたかも心の状態であるかのように受け入れてしまう」。このように「心の状態」を社会的に構成し承認することで、責任追及を可能にするのである。

「過失とはじつは心の状態ではなく、あらかじめ気がつくべき法益（歩行者）が決まっており、その法益侵害をしたときに（故意でなければ）自動的に『気がつかなかった』とみなされる社会制度すぎない」。それを社会が「合理的」なものとして容認しているにすぎない。それゆえ、「事故が生じたときに法律を適用するさいの処理」は、一方では「説得的」であるが、他方で「虚構的」でもあるということになる。<sup>(40)</sup>

## おわりに

西垣通は、「自由意思（free will）」をもっていることが人間が「道徳的主体」であるための「大前提」であるという。<sup>(41)</sup> 自由意思をもつことによって、人間は道徳的主体として判断をくだし行為を選択する。そして「その結果に『責任（responsibility）』をとらなくてはならない」。では自由意思の有無はどのように確認すればよいか。それは「理論的自律性（theoretical autonomy）」が認められるかどうかによってである。理論的自律性は、下等な生物でも持っているような特性である。生物はオートポイエティック（自己-創出的）な存在であり、ある環境条件のもとで実行される行為の内部ルールを自分で作りあげる。環境条件の変化により内部ルールは変化することもあるが、どのようなルールが出現するかは、「当の生物をふくめ誰にもわからない」。こ

の「不可知性（絶対的な予測困難性）」が、生物の場合には成立している。そのため生物は「他者の指令をまったく受けずに行動する」。これが「理論的自律性」である。これと区別されるのが「実践的自律性（practical autonomy）」である。これは、道徳的判断をくだし「社会的責任をとれる主体のもつ特性」であるとされる。AIの場合、どのような入力に対してどのような出力（行動）が現れるのかが、設計者には基本的に分かっている。細部は不明でも、原則として行動が予測可能である。AIは、生物とはちがひ、人間が設計するアロポイエテック（他者-創出的）な存在であり、予測困難性が成立しない。それゆえ、AIには理論的自律性も自由意思もなく、責任を問うこともできないことになるという。

だが、これまで述べてきたように、責任を負うべき結果が生じ責任が問われるときに責任追及の終止点として事後的に「構成」されるのが自由な意志であるとすれば、AIに関しても「自由意志のようなもの」を「構成」することは不可能ではないだろう。問題は、そうした新たな「虚構」としての「約束ごと（nomos）」を社会が合理的なものとして容認できるか否かである。

稲谷龍彦は、人と機械の〈協調動作〉が日常に組み込まれることで、結果的に意思決定や行為の主体を特定することが困難になるという<sup>(42)</sup>。そのため、AIを搭載している機械が人と協調動作している際に何らかの事故が生じた場合、誰が法的責任を問われるのかという解答困難な法的問題が生じる。従来刑事法は、主体の自由意志に基づく行為に対して刑事制裁を科し、社会秩序を維持するという構図を採用してきたが、今や、自由意志を持つ主体と主体の影響下に置かれる客体という区別は消滅しつつある。刑事法理論は、人間という存在を所与の前提としてきたが、この前提の限界が露呈しつつあるのだという。そして、人間がもはや「世界の所与の構成単位」とはいえないという主張を「ポスト・ヒューマニズム」と呼んでいる。行為や行為者性といった概念もまた、これまでは人間を所与の前提として考えられてきたとあってよいだろう。そしてポスト・ヒューマニズムにおける倫理では、「人

問」を前提とせずにこれらの概念について考える必要があるだろう。

〔注〕

- (1) 佐久間修「AIの刑事責任」刑法雑誌第59巻第2号(2020年)、今井猛嘉「AI時代の刑事司法」罪と罰第56巻第2号(2019年)、川口浩一「ロボットの刑事責任2.0」刑事法ジャーナル第57号(2018年)、今井猛嘉「自動車の自動運転と刑事実体法」山口厚・佐伯仁志・今井猛嘉・橋爪隆(編)『西田典之先生献呈論文集』(有斐閣・2017年)。
- (2) 今井猛嘉「自動走行に関与する者の刑事責任」NBL(New Business Law)1099号(2017年)、25-29頁。
- (3) 古川伸彦「自動運転車、死傷事故、刑事過失責任」刑法雑誌第59巻第2号(2020年)、319頁。
- (4) 深町晋也「自動運転車に関するAIを巡る刑事責任について」罪と罰第56巻第2号(2019年)、43頁以下。
- (5) 中川由賀「自動運転導入後の交通事故の法的責任の変容」中京LAWYER Vol.25(2016年)、50頁。
- (6) 根津洗希「ロボットの処罰可能性を巡る議論の現状について」比較法雑誌第51巻第2号(2017年)、147頁。
- (7) 池田良彦「自動運転走行システムと刑事法の関係」自動車技術第69巻第12号(2015年12月号)、35頁。
- (8) 中山幸二「自動運転をめぐる法的課題」自動車技術第69巻第12号(2015年12月号)、41頁。
- (9) 橋本祐子「刑事司法における『感情』の所在」論究ジュリスト第22号(2017年)、34-40頁。
- (10) 戸田正直『感情：人を動かしている適応プログラム』(東京大学出版会・1992年)、1-5頁。
- (11) 同書、10-22頁。
- (12) Lerner, M.J., *The Belief in a Just World: A Fundamental Delusion*, (Springer, 1980)。
- (13) 大淵憲一『謝罪の研究：釈明の心理とはたらき』(東北大学出版会・2010年)、58-59頁。
- (14) 同書、18-26頁。
- (15) 同書、4頁。
- (16) 同書、5-18頁。



- (17) 同書, 65-67頁。「誤解を与えたとすれば申し訳ない」といった釈明の言葉が「フェイク謝罪」と呼ばれている。これは、言語表現として「悔悛表明」を含むが真正な謝罪ではない。非難の対象を「発言者の考え(発言内容)」から「表現の稚拙さ」にすり替えており、受け手側が「誤った理解」をしているという指摘によって、発言者(加害者)ではなく受け手(被害者)の側に「誤り」があると主張している点で、「責任転嫁」と「加害の上塗り」にすらなっているとされる(「論の芽」朝日新聞2023年5月31日朝刊)。発言者の発言内容が受け手に精神的苦痛という被害をもたらしたのだとすれば、本来であれば、発言者の「謝罪」は、加害行為への関与とその不当性を認め、自分が責任を負うことの表明でなければならない。「私の発言により精神的苦痛を与えてしまい申し訳ありません」といった謝罪の言葉が発せられるはずである。しかし、「発言が誤解を与えたとすれば申し訳ありません」と言い換えることにより、その意味はいくつかの点で大きく変わる。受け手の精神的苦痛に関与しているのが本人の「誤った理解」であるとすれば、その「生成責任」は受け手にあることになるから、発言者が非難対象となることは否定される。「誤解を与えたとすれば」という表現は、その誤解の背景として発言者の表現技能の未熟さがあった可能性を述べているに過ぎない。「謝る」は「誤る」から派生した語であり、誤りを認めて許しを請うところから、「あやまる」に「わびる」「謝罪する」の意味が生じた。「あやまる」が謝罪の意味で一般に使われるようになったのは、中世末から近世にかけてであるとされ、これに「謝」の漢字が当てられるようになったのは近代以降とされる。受け手の理解が「誤っている」とすれば、本来「謝る」べきは受け手である。受け手の精神的苦痛は、自身の誤解が直接的な原因となってもたらされたものということになる。発言者には「生成責任」「有責責任」もなく、当然「負担責任」を負うこともないが、受け手の「誤解」をもたらした一因(間接的な原因)として自らの表現技能の未熟さがあったとすれば、それについては「謝罪」と言っているに過ぎない。その意味で「誤解を与えたとすれば申し訳ない」という釈明は、受け手に苦痛をもたらしたことに対する「謝罪」ではなく、自らの発言に「生成責任」を認めていない点では「否認」に分類されるものである。しかしながら、表面的には「謝罪」したようにみえるという意味で「フェイク謝罪」である。実態としては「責任受容」なき「悔悛表明」といえよう。
- (18) ヴィンフリート・ハッセマー『刑罰はなぜ必要か：最終弁論』(中央大学出版部・2012年), 5頁。
- (19) 『新明解国語辞典第八版』(三省堂・2022年)。
- (20) 『日本語源広辞典 [増補版]』(ミネルヴァ書房・2012年)。

- (21) 瀧川裕英『責任の意味と制度：負担から応答へ』（勁草書房・2003年），16-17頁.
- (22) Hart, H.L.A., *Punishment and Responsibility*, (Oxford UP, 1968), p.211.
- (23) 瀧川・前掲注 (21), 38頁.
- (24) 大淵・前掲注 (13), 63-64頁.
- (25) 同書, 51-55頁.
- (26) 大塚裕史・十河太郎・塩谷毅・豊田兼彦『基本刑法 I : 総論 [第3版]』（日本評論社・2019年），25頁.
- (27) 池上俊一『動物裁判：西欧中世・正義のコスモス』（講談社現代新書・1990年），26頁.
- (28) 大屋雄裕「AIにおける可謬性と可傷性」宇佐美誠（編）『AIで変わる法と社会』（岩波書店・2020年），45-46頁.
- (29) エドワード・ベイソン・エヴァンズ『殺人罪で死刑になった豚：動物裁判にみる中世史』（青弓社・1995年），181-182頁.
- (30) 池上・前掲注 (27), 180-215頁.
- (31) Kelsen, H., *What is Justice?*, (University of California Press, 1957), pp. 329-330.
- (32) 小坂井敏晶『増補責任という虚構』（ちくま学芸文庫・2020年），231頁.
- (33) 黒田亘『行為と規範』（勁草書房・1992年），69頁.
- (34) 同書, 65-68頁.
- (35) アリストテレス『ニコマコス倫理学』（『アリストテレス全集13』（岩波書店・1973年）），第3巻第1章.
- (36) 野家啓一「原因／理由」『岩波哲学・思想事典』（岩波書店・1998年），442頁.
- (37) Wittgenstein, L., *Philosophische Untersuchungen*, (Blackwell, 1953), §621.
- (38) 黒田・前掲注 (33), 9-11頁.
- (39) 中島義道『時間と自由：カント解釈の冒険』（講談社学術文庫・1999年），138-163頁.
- (40) 中島義道『後悔と自責の哲学』（河出書房新社・2006年），26-29頁，41-43頁. 小坂井敏明も，そうした「虚構性」を明らかにしている. 小坂井・前掲注 (32).
- (41) 西垣通・河島茂生『AI倫理：人工知能は「責任」をとれるのか』（中央公論新社・2019年）58-67頁.
- (42) 稲谷龍彦「ポスト・ヒューマニズムにおける刑事責任」宇佐美誠（編）

『AIで変わる法と社会』（岩波書店・2020年），113-136頁．

**【付記】**

本研究は JSPS 科研費 JP21K00015 の助成を受けたものである．