ザビーネ・グレス＝トーマス・ヴァイгент
「インテリジェント・エージェントと刑法」
Sabine Gleß/Thomas Weigend, Intelligente
Agenten und das Strafrecht

伊藤嘉亮

- I. 導入
- II. 文献紹介
 - 1. IA——ポテンシャルとリスク
 - 2. IAの主体性
 - 3. IAの背後者
 - 4. 結語
- III. コメント
 - 1. IAの可罰性
 - 2. IAの背後者の可罰性

I. 導入

広義のエージェントは、①人間エージェント、②ソフトウェア・エージェント、③ハードウェア・エージェントに分類される。①は、特許や実用新案などの申請・出願などの代理を職業とする特許弁理士、客の要望に応じて航空機やホテルの予約をしてくれる旅行案内業者など、人間が動作主体となるものである。これに対して、コンピュータ・ソフトウェア(例えば、情報収集エージェントなど)が②であり、ロボット(例えば、掃除ロボットなど)のように特別なハードウェアを伴う場合は③となる⁽¹⁾。

②・③のうちでもインテリジェント・エージェント(以下、IA)⁽²⁾と称

(1) 荒屋真二『人工知能概論——コンピュータ知能からWeb知能まで——〔第2版〕』(共立出版、2004年)143頁以下。

されるものは、「自らの知識や問題解決・意思決定の仕組みを持ち、おかれた環境からの作用に対応して自律的に行動するという特徴を持っている」⁽³⁾。つまり、IAは、その機能として、(1)環境の認識(「環境認識機構」が周囲の状況(環境)に関する情報をセンサーを通じて受け取り、「ルールベース」を活用して環境を認識・解釈する)、(2)探索・推論(環境認識機構からの情報に基づき「推論機構」が目標を設定し、その目標達成に必要な知識を探索し、推論計算によって問題解決を行う)、(3)環境の操作(推論機構による問題解決の結果に基づいて「環境操作機構」が具体的な操作を決定し、環境へ働きかける)、(4)学習(「学習機構」が過去の行動の経験に基づいてルールの価値の変更やルールの追加・削除を行い、より迅速かつ適切な行動をとれるよう学習する)についての能力を有するのである⁽⁴⁾。

今回紹介する論文⁽⁵⁾は、以上のように認識・決定・行動のすべてを自律的に行いうるIA自体の刑事責任、およびIAを開発・利用する周囲の自然人の刑事責任を考察するものである。前者の問題が現実味を帯びるのはもうしばらく先の時代の話であろう。しかし、他方で、2020年には自動運転車が我が国でも実用化されるのであれば、自動車メーカーや自動運転車の利用者の刑事責任を問う後者の問題は、喫緊の課題として取り上げられるべきである⁽⁶⁾。ここでの議論の中心を占めるのは開発者や利用者の過失責任の成否、とりわけ注意義務や予見可能性をめぐる問題であると思われるが、IAという先端技術は従来の過失犯論にその修正を迫るものなのだろうか。本論文は、この問題を考える上での重要な視

(2) 一般的には「知的エージェント」と呼ばれている。

(3) 前田隆＝青木文夫『新しい人口知能〔発展編〕』(オーム社、2000年)172頁以下。

(4) 荒屋・前掲注(1)146頁以下。

(5) Sabine Gleß/Thomas Weigend, Intelligente Agenten und das Strafrecht, ZStW 126(2014), S. 561 ff.

(6) この問題については、今井猛嘉「自動化運転を巡る法的諸問題」国際交通安全学会誌40巻2号(2015年)132頁以下、中川由賀「自動運転導入後の交通事故の法的責任の変容——刑事責任と民事責任のあり方の違い——」CHUKYO LAWYER 25号(2016年)41頁以下参照。

座を提供してくれるものと思われる。

II. 文献紹介

1. IA——ポテンシャルとリスク

IAは、「グーグルボット」や「グーグルオートコンプリート」のように事前に設定された一定のルールに従って作動するが、個別事例においては情報(グーグルのユーザーの検索)を独立に処理するものであって、インターネットの検索エンジンとしてだけでなく、オンライン上の株取引において自動で売買を決定するソフトウェア・エージェントとして、または軍事的、プライベート空間の監視・保安、あるいは商品の配送に用いられるドローンとして我々の生活の多くの場面で見受けられる。また、現在の我々が最も鮮明にイメージするIAは、運転手のいない自動運転車であると思われる。

IAは、膨大な情報を取得するだけでなく、一定のパターンに従ってそれを解釈し、そうして処理された情報に対して独立に〔人間の指示を必要とすることなく〕反応するものであるが、更に、将来のためにそれらを記憶することができる。また、所定の目的を可能な限り効率的に実現するために、情報を整理するだけでなく、学習することもできる。自律的に機能する自動運転車の場合、自動運転車はセンサーを通じて得た情報をパターンに基づいて検討して直接それに反応するものであるから、第三者からすれば自動車自体が行動しているようにも感じられる。更に、開発者すらも特定の交通状況における当該自動車がどのように作動するのか確かなことは出来ない。それゆえ、自動運転車のエンジニアは、事故が何件が発生した後にはじめて、例えばライトや天候が特定の状況にある場合に当該自動車が優先通行標識を認識しなくなることを知りうるのである。

IAは、想定外の新たな状況に対応する必要がある場合でも、人間が与えた任務を遂行しようとするものであり、ここにIAのベネフィット(Nutzen)が認められる。しかし、それが故に、人間は、そのような想定外の状況下に置かれたIAの動作を予見することも制御することも出来

ない。学習可能なロボットの価値と制御可能性の欠如は表裏の関係にあるのである。

学習可能なIAは人間を新たな次元へと進歩させるものであるが、それと同時に、IAにより刑法上重要な結果(例えば、人の死傷や財物の損壊)が惹起された場合に誰が責任を負うのかという難問をもたらす。刑事責任を負うのは、当該テクノロジーを起動させた人間であろうか。それともIA自体であろうか。あるいは、そうしたテクノロジーがもたらすベネフィットを欲する社会こそが当該テクノロジーの予測不可能性により生じる結果を受け入れるべきであって、刑法上は誰も責任を負わないのであろうか。

2. IAの主体性

(1) 前例〔省略〕

(2) IAの人格性

ロボットが周囲の環境から得たデータを既知のパターンと比較し、それを基に行為する能力を予めプログラミングされていることをもって、当該ロボットをその態度につき責任を問われる「人格」とみなすのに十分といえるだろうか。

カントの法哲学によれば、自己反省能力が人格に自己決定を可能ならしめるのであって、また、人格が自由の自意識において自身の所為に責任があると考えられる場合に、他者も当該人格にその責任を問う権限を持つことになる。これに対して、IAには人間性(Person-Sein)の本質的要素が欠如している。たとえIAが学習でき、他者が予見できない決定を下しようとしても、自分自身の自由を意識することはできないし、また、自らを社会における権利・義務の所有者と考えることもない。IAは、確かに一定の任務を「自動的に」処理できるが、結局はプログラミングを通じて予め与えられた選択肢に応じるのであって、自らの態度について自らに釈明することはない。それゆえに、我々は、IAを「不自由」と感じるのであり、それらが引き起こした損害につき、それら自体に対して人格的に責任を問うことはしないのである。

(3) IAの可罰性

(i) IAの行為能力

インターネットの検索エンジンが名誉を毀損する組み合わせで単語を併記し、ドローンが無辜の人を害し、運転手のいない自動運転車が歩行者を轢く。これらは、IA自体の行為であるようにも思われるが、本当に刑法上の行為といえようか。

因果的行為論においては、刑法上の行為としてはあらゆる任意の身体動作で十分である(因果的行為論の現代版である社会的行為論は更に何らかの社会関係を要求する)。IAが「任意に」行為しうるかも問題になるが、因果的行為論の立場からは、意思制御に対してはそれほど高度なものが要求されるわけではなく、(制御できない反射の場合のように)単に「自動的に」反応した場合や他の人間により物理的な対象物として利用された場合に任意性は否定されるに過ぎない。これに対して、目的的行為論によれば、人間の行為は、行為者の目的志向的な意思を表現する場合にのみ、かつ、それが故に刑法上重要になりうるわけだが、IAは自ら「意図的に」目標を設定し、その目標設定に従って行為を定めることはできない。

結局のところ、IAの行為能力は定義次第ということになる。あらゆる任意の身体動作と解するならばIAを行為者とみなしうるが、行為概念の中に自意識による目標設定を読み込むのであれば、IAが行為能力を有することはない。

(ii) IAの責任能力

BGHは、責任について、「責任非難の内在的根拠は、人は道徳的に成熟するや否や、自由かつ道徳的に自己決定する能力が病的事象により一時的に鈍麻していない、あるいは長期間阻害されていない限り、自由、答責的、かつ道徳的に自己決定する能力を有し、それゆえ適法に従い、不法に抗う決定をし、彼の態度を法的当為の規範に適合させ、法的に禁止されたことを回避することができる、という点にある」と定義した。そうすると、IAは道徳的に自己決定する能力を有する人格ではなく、与えられた任務をプログラムに従って実行するコンピューター・テクノロジーであると考えられていることから、IAの責任は排斥されるよう

に思われる。

他方で、刑法上の責任を認定するにあたっては現実の社会関係における期待と帰責(Zuschreibung)が重要であると解する場合、責任は紛争の答責的な発起者(Urheber)と同定される者に帰せられるのであって、また、自由意思といった概念も一定の社会的な目的のためになされる帰責の産物であって、如何ようにでも解釈されうるものになる。しかし、責任は、〔何らの制限もなく〕任意に帰せられるものではない。自分自身および過去の自分の決定を倫理的な評価基準体系により評価できない者、つまり「善悪の判断」のつかない者とは倫理的な対話は不可能であって、彼が倫理的な非難に対して「返答」することもできないため、責任の条件としては自己反省能力が要求されることになる。

IAに実装される規範的なルールの遵守も考慮したボーナス・マルス・システム〔業績に応じて報酬等を増減するシステム〕が機能的に人間の倫理的な自己評価・自己統制に相当するのであれば、IAに帰責することも可能になる。しかし、倫理的な制御システムの実装はプロジェクトの端緒についたばかりであり、ロボットにこれに基づく決定をプログラミングするのはまだまだ先のことといえる。

(iii) IAの処罰

IAは自身の財産を所有しないため、それに科される罰金刑は〔結局のところ〕その操縦者や負責基金により返済されるのであって、IAには向けられない。IAの物理的な破壊やプログラムの書き換えによる「人格的」処罰は一種の死刑のようにも思われるが、IAが人間と同様に長く生存しようとする意思を有していない限り、人間に対する制裁と同じ態様でIAに向けられるわけではない。また、人間はロボットの破壊をその過誤に対する非難的制裁として理解できるが、当該処分のそうした意味がロボット自体に伝わることもない。

もっとも、IAに(不完全であったとしても)刑法上の答責性を認めることには、それが背後にいる自然人の責任を否定しうる限りで法的に意味がある。

(iv) 小括〔省略〕

3. IAの背後者

IAに独立の刑事責任を認め得ない以上は、IAの過誤についてその背後者に責任を問いうるかが問題となる。

(1) 故意処罰

事業主が刑法上重要な結果を惹起するよう計画的あるいは故意にIAをプログラミングしたのであれば、当該事業主に刑事責任を認めることに問題はない。軍の指揮官が私人を殺害するために意図的にドローンを使用した場合、あるいは設計者が被害者を轢くよう故意に自動運転車をプログラミングした場合、彼らは正犯として責任を問われる。というのも、彼らは、その他の事例で道具を思い通りに使用するのと同様にIAを用いているからである。

(2) 過失処罰

IAの使用が意図しない損害を惹起する事例の方が数としては多い。ドローンは、事業主の意思に従って事前に与えられた軍事目的のみを追求するものであるが、機能不全によりそのルートから外れ、一般市民を殺害することもある。自動運転車は、通常、他者の優先通行を許可するものであるが、これを無視して他者を害することもある。これらの場合に事業主が構成要件の結果を予見でき、当該結果を回避するために必要な注意を払わず、更に結果が事業主に帰属されるのであれば、彼は過失犯として処罰されることになる。

(i) 予見可能性

IAは、周囲の状況から取得した情報を独自に評価し、人間からの指示を必要とすることなくその任務を遂行するものであるため、結果の予見可能性の時点で既に問題が生じる。つまり、事業主はIAがどのように作動するのかを予見できないのであって、それゆえ、IAの「一定程度の予測不可能性」およびそれに伴う「第三者に対する危険」は前もってプログラムに組み込まれているのである。

この点、独立の情報処理に基づく機械の侵害行動は予見し得ないのだから、自律的に作動するエージェントの事業主を過失犯としては処罰できないとの見解もあり得るが、これは妥当でない。例えば動物園の責任者が虎を解放し、その虎が通行人に傷害を加えた場合に、野生動物の予

測不可能性を理由に彼を不処罰とするわけでないのと同様である。それを理由に不処罰とするならば、自然人とIAの間で答責の拡散(Verantwortungsdiffusion)をもたらし、IAによる被害を放置することになる。したがって、考えられうるあらゆる損害結果は事業主には予見しうるものとせざるを得ず、そうであれば予見可能性という要件はもはや事業主に過失責任を認める上での障壁ではないことになる。

(ii) 注意義務違反

過失処罰は、構成要件実現の予見可能性と並んで注意義務違反を前提にする。

潜在的に危険な商品は、その安全性が技術水準に合致し、販売の前に十分なテストにより検査された場合にのみ市場に出すことが許される。製造者は、商品を流通させて以降も消費者のレスポンスに基づいて引き続き監視しなければならない、その際に予期せぬ損害やリスクが示された場合、製造者には消費者への警告が義務づけられ、必要な場合には製品の回収を義務づけられる。こういった原則は、IAのような革新的な製品に関しても妥当する。刑事処罰の可能性があることにより、IAはその制御可能性の欠如ゆえに製造・販売されなくなるであろう。確かに、こうした帰結は安全性の確保にも結びつくが、他方で、我々の生活の負担を著しく軽減する多くの革新的技術を喪失することにもなる。

しかし、過失犯として処罰される可能性があるからといって、リスクを伴うあらゆる行為を完全に放棄しなければならないわけではない。革新技術からベネフィットを得る「社会」がそれに伴うリスクを事業主に負担させるのは公平でないから、事業主が創出するリスクを当該テクノロジーのベネフィットと関係づけながら注意義務を定める必要がある。ここでは、総効用(Gesamtnutzen)の増大を理由に社会がIAに伴う危険をどの程度受け入れるかが重要となる。IAの使用に伴うリスク——例えば、インターネットの検索エンジンに記録されたデータの不適当な使用——が既に一般的に受け入れられている場合、IAが(一般的に)予見可能な損害をもたらしたとしても、事業主の刑事責任は否定されるのである。

一方で、IAのベネフィットがどれ程高く、他方で、IAが生じさせうる危険がどれ程重いかが重要となる。IAが有するベネフィット次第で

は、その開発・利用に伴うリスクを甘受することになろう。そうすると、十分なテストを行なったのであれば、計画外の状況に遭遇した場合に例えば突然停止することが希にあるとしても、自動運転車を流通させることは許容される。自動運転車は個人にとっても社会にとっても大きなポテンシャルを持つものであるから、社会的な付加価値の少ないIA(例えば玩具や戦闘用ドローン)に適用されるものより寛大な基準が妥当しうるのである。もっとも、保護法益のための注意義務は、有益なIAの場合にもリスクを最大限排除することを要求する。それゆえ、革新的技術のリスクが受け入れられている場合であっても、例えば、ユーザーに対する包括的な警告・指示義務、あるいは製品の使用状況(Erfahrungen)を観察し、損害発生の報告に対して折り返し対応する義務はある。例えば、自動運転車による侵害が極めて希な状況や著しい誤用の場合のみに限られないのであれば、当該自動車をそれ以後も販売することは許されず、場合によっては回収しなければならないのであって、それにもかかわらず事業主が何らの行動にも出ず、損害が発生したならば、処罰されうるのである。

(iii) 客観的帰属

IAの事業主が注意義務に違反したとしても、更に、構成要件の結果が彼に「客観的に」帰属されうるか否かを問うことになる。

第一に、IAが惹起した結果(例えば、自動運転車による人の傷害)は、人間によって統制されたわけではない自然現象なのであって、そこでは単に一般的な生活リスクが実現したに過ぎないのであれば、事業主への当該結果の帰属は否定される。これは、しかし、IAが日常生活の中で一般的に存在するものと考えられていることを前提にするのであって、IAが未だ例外的な存在である現在および近い将来においては、それを流通させる者は異常なリスクを創出している。もっとも、既に検索エンジンやレコメンダ・システムといったソフトウェア・エージェントの場合、我々は自身的人格権が侵害される危険を認識しているにもかかわらずオンライン・ショップでの購入を継続するのであって、それを一般的な生活リスクとして受容している。また、運転手のいない地下鉄や自動車が鉄道・道路交通における当然の光景になったのであれば、そのよう

な乗り物の利用は人間にとって通常のリスクとなり、事業主は、回避可能な設計またはプログラミング上のミスに基づく侵害についてのみ責任を負うことになる。

第二に、自律的に作動するロボットの「自己」決定が事業主の行為と結果との間の帰属関係を遮断する場合がある。確かに、IAに対する刑事制裁が問題になることはない。しかし、IAの動作がその自律的な行為と認定されうる場合、背後者(事業主)への帰属関係は否定されるのである。もっとも、IAに意思を自ら形成する能力を認めない限り、IAの侵害行為が自律的であるとはいえないが、現在のところそれは不可能である。また、たとえIAがその域に到達したとしても、IA自体の処罰が認められないのであれば、過失の認められる事業主への遡及を許容すべきであろう。さもなければ、人間とIAの間での答責の拡散が生じてしまうからである。

4. 結語

人間はIAを完全にはコントロールできない。それにもかかわらずIA〔の開発・使用〕を全面的には禁止しないのであれば、事業主の過失責任を否定し、IAに伴う危険の責任を社会に負わさざるを得ない。ただし、IAの背後者の責任が否定されるのは、IAのベネフィットがそのリスクの受容を正当化する場合に限られる。

以上の帰結は、例えば「開発リスク(Entwicklungsrisiko)」の許容を通じて事業主に要求される注意義務を限定することによって、あるいはIAの「社会的に相当な使用」という正当化事由を創設することによって導かれる。もっとも、法律の明確性という観点からは許されるリスクの基準をIAの種類・分野ごとに法律または命令で定めるべきである。

Ⅲ. コメント

本稿は、自律的に周囲の状況を把握・分析し、人間の指示を必要とすることなく作動するIAをめぐる法的問題を論じるものであり、IA自体の可罰性とIAの背後者たる自然人の可罰性の両者を検討している。前

者については、IAの行為能力、責任能力および受刑能力を考察し、現時点でのIAの可罰性を否定している。後者については、背後者が故意にIAを利用する場合には問題なくその罪責を認めうとする一方、過失犯の成否が問われる場合(つまり、意図しない侵害結果が発生した場合)には、IAという存在が予見可能性、注意義務および客観的帰属のいずれにも影響を与えることを示しつつ、IAのリスクとベネフィットを中心に議論を進めている。以下ではそれぞれの問題について、若干の考察を加える。

1. IAの可罰性

グレス=ヴァイгентは、前述のように、因果的行為論の立場からはIAの行為を観念することはできると主張している。確かに、この立場によれば、自動運転車やドローンのように物理的なボディーを有するIAの行為を観念することは容易であろう。これに対して、インターネット上の検索エンジンのような純粋なソフトウェア・エージェントの場合、人間の「身体」に相当するものがないため、その行為能力を肯定しうるかが問題となる。この点、行為論において「外界への影響」を重視し、また、それのみをもって行為能力の十分条件と解するならば、ボディーを伴わないIAの行為を肯定することもできよう。他方で、「身体の動静」を行為の必要条件として要求するならば、IAの可罰性は、既に行為能力の時点で否定されるように思われる。

次に、グレス=ヴァイгентは、責任の認定が所与の事実の発見ではなく、犯罪を誰かに帰責するための作業であると解した上、そうした帰責の前提として善悪の判断力、倫理的非難に対する返答可能性、およびそれらに基づく倫理的対話能力を要求し、(少なくとも現在の)IAがそれらを備えていないことから、IAの責任能力を否定している。確かに、責任評価を「評価側(裁判官)による非難」と「被評価側(行為者)による了解・納得」の相互応答と捉えるならば、IAに責任を認めるには、被評価側であるIAがその行為の社会的意味を理解・評価できなければならないことになる⁽⁷⁾。しかし、今のところ、そこまでの機能を備えたIAは開発されていないため、IAの責任能力は否定されざるを得ないだろ

う。また、受刑能力においても、刑罰を法的非難の伝達手段と考える⁽⁸⁾のであれば、刑罰の非難の意味を理解し得ることが受刑能力の必要条件となろう。しかし、現在のIAがこの要件を充足することはないと思われる。

結局のところ、グレス=ヴァイгентが主張するように、現在の開発状況を前提にする限りはIA自体を刑法の行為主体と解することは不可能である。他方で、グレス=ヴァイгентが、IAの責任・受刑能力を否定しつつも、その答責性を認めることには周囲の自然人の刑事責任を検討する際に意義があるとしている点は興味深い。つまり、IAを——刑事責任を有さない点で共通しつつも——他の自然物や製造物からは区別し、犯罪事実におけるアクターの一人として捉えることにより、IAは、これを使用する自然人の刑事責任を否定する一要素になりうるのである⁽⁹⁾。

2. IAの背後者の可罰性

(1) 故意処罰

IAの背後者の故意犯を論じるにあたってグレス=ヴァイгентが念頭に置くのは、背後者が①違法な行為を行うよう予めプログラミングさ

(7) アルトゥール・カウフマン(上田健二訳)『法哲学〔第2版〕』(ミネルヴァ書房、2006年)170頁以下は、「ある犯罪行為の意味は、あたかも先置きされたもの、固定したものであり、帰責に当たってはこの意味を行為者の表象のなかに再発見することだけが問題であるかのように考えられている」が、(刑法上の)帰責は、「行為者と裁判官との間におけるひとつのコミュニケーション過程として特徴づけ」られるとする。

(8) こうした理解に対しては、松原芳博「刑罰の正当化根拠としての応報——刑法学の視点から——」応報の行方(法哲学年報2015年)70頁が、「刑罰と非難との関係は、刑罰が非難を含むという前提から応報刑の正当性が導かれる一方で、刑罰が応報であるという前提から非難の契機の必要性が導かれている点で循環論法に陥っている」と指摘する。

(9) スザンネ・ベック(只木誠=富川雅満訳)「ロボット工学と法——その問題、現在の議論、第一の解決の糸口——」比較法雑誌50巻2号(2016年)110頁以下の「電子的人」概念も参照。

れたIAを②意図的に利用する事例である。このような場合であれば、確かに、IAのその後の動作を詳細には予見できないとしても、動物を利用して他者を侵害した者を正犯として処罰しうるように、IAの背後者を正犯として処罰することに争いはないと思われる。

しかし、背後者の故意犯が問題になる事例はそのような場合には限定されない。上記①とは異なり、法令を遵守するようプログラミングされたIAを故意に用いる場合も想定できるし(①)、また、確定的故意をもってIAを利用する場合(②)のみならず、プログラムに重大なバグがある可能性を未必然に認識しながらそのIAを販売・利用する場合も考えうる(②)。

①の場合、たとえ背後者が当該IAに違法行為を実行するよう命令したとしても、法令を遵守するようプログラミングされているため、その命令を遂行する可能性はないとして、当該命令行為を不能犯と解する余地もある。また、当該IAは適法行為に出るよう期待される、あるいは違法行為を遂行するか否かを自律的に決定しうるのであるから、規範的障害説⁽¹⁰⁾または遡及禁止論⁽¹¹⁾の立場から背後者の正犯性を否定することもできよう⁽¹²⁾。②については、バグの性質(当該バグにより危殆化される法益の種類や範囲、程度、またはそうした危険が現実化する蓋然性)次第では、バグを放置したまま販売・利用したとしても、許された危険の範囲内にあるとして背後者の可罰性を否定することも可能であると思われる。

(2) 過失処罰

(i) 予見可能性の拡張の必要性

IAの背後者に過失責任を問う上で最初に問題になるのは、予見可能

(10) 曾根威彦『刑法原論』(成文堂、2016年)527頁以下。

(11) 島田聡一郎「間接正犯と共同正犯」『神山敏雄先生古稀祝賀論文集第一巻』(成文堂、2006年)447頁。

(12) 他方で、IA自体の刑事責任が否定される根拠を「責任の欠如」に求めるのであれば、IAを正犯とする教唆犯・幫助犯を観念しうるかもしれない。もっとも、刑法61条1項が規定する「人」および同62条1項が規定する「正犯」がいずれも自然人を意味するのであれば、IAに対する教唆犯・幫助犯は否定されることになろう。

性の有無である。この点、グレス＝ヴァイゲントは「考えられうるあらゆる損害結果は事業主にとって予見可能なものとせざるを得」ないと主張しているが、これが、従来の「予見可能性」概念を維持したままでもその存在を肯定し得るとの主張なのか(①)、予見可能性を従来の理解よりも幾分拡張すべきとするものなのか(②)、はたまた我が国における危惧感説と同様に過失犯処罰には「危惧感」で足りるとする趣旨なのか(③)は明らかでない。

例えば動物による他者侵害について過失犯が問われる場合、動物も自ら周囲の状況を確認しながら行動するのであって、その行動を詳細に予見することはできないが、それでも当該動物の背後者(飼い主)に予見可能性を肯定し、過失責任を認めうることからすると、IAを利用する場合であっても、その運用上想定できる侵害結果の枠内においては、たとえ侵害に至るプロセスを具体的に予見することはできなかつたとしても、当該侵害結果の予見可能性を肯定する余地はある⁽¹³⁾。したがって、IAが介在する過失事例を適切に解決すべきことは当然の要請であるが、この要請は、「予見可能性」概念を②または③のように拡張すべきとする契機にはならない。また、IAの開発・利用の時点において、背後者にはおよそ想定不可能な原因(プログラム上のバグ等)に基づき侵害結果が発生した場合であれば、やはり当該結果の予見可能性、ひいては過失犯の成立は否定されるべきであって、「予見可能性」概念を②または③のように拡張して背後者の予見可能性を肯定することは、責任主義の観点からは是認できないと思われる。

(ii) 実行行為の成立範囲の限定

予見可能なあらゆる結果について背後者(開発者や利用者)に過失責任

(13) 刑事製造物責任における予見可能性については、例えば、甲斐克則「欠陥製品の製造・販売と刑事過失」『神山敏雄先生古稀祝賀論文集第一巻』(成文堂、2006年)169頁以下を参照。IAの開発や販売も刑事製造物責任の一種といえるため、刑事製造物責任をめぐる従来の議論は参考になるとと思われる。また、判例が用いる「因果関係の基本的部分」公式については、杉本一敏「『因果関係の基本的部分』の予見可能性について——渋谷温泉爆発事件決定を契機に——」『刑事法ジャーナル50号(2016年)9頁以下を参照。

を問うのであればIAなどの技術革新を阻害しかねないため、背後者の過失責任を適切な範囲に限定しようよう要請される。グレス＝ヴァイゲントが指摘する「IAの開発・利用に伴うリスク」と「IAが社会全体にもたらすベネフィット」が背後者の過失責任を画する際に考慮される要素であることは間違いないが、本論文は、これらと注意義務の関係や過失責任の具体的な認定プロセスを明らかにしていない。そこで、以下では、過失犯の注意義務や実行行為の具体的な認定プロセスについて若干の検討を行う。

①まず、開発者らに対してはIAの開発・利用に伴うリスクを減少させるよう注意義務が課される。そこで課される注意義務の具体的内容は「リスク」と「コスト」を考慮しながら決せられることになろう。つまり、比例原則の観点から、リスクを同程度に減少させる注意義務が複数考えられるのであれば、コストの低い注意義務が選択されることになる。また、例えば、あるIAが創出するリスクが潜在的被害者の生命に関係するのであれば、コストの相当高い注意義務を課することも正当化され、他方で、そのリスクが軽微な利益に関係するに過ぎないのであれば、コストの高い注意義務の賦課は否定されよう⁽¹⁴⁾。②そのようにして選択された注意義務を適切に履行することによってリスクは確かに減少するが、リスクが完全にゼロになることはほとんどないだろう。そこで、次に、注意義務を履行しながらも残存した「リスク」とIAがもたらす「ベネフィット」を比較衡量し、その結果として社会が当該リスクを受容するかが検討されるべきである。③比較衡量の結果として社会が受け入れたリスクは「許された危険」であり、過失犯の実行行為性が否定されることから、たとえそのリスクが実際に実現したとしても背後者に過失犯が成立することはない⁽¹⁵⁾。他方で、注意義務を履行して減少さ

(14) 「注意義務の設定時点における危険の内実が明らかにされていることを前提に、①当該危険防止の有効性、②当該危険に比して義務内容が過大でないこと、③複数の措置が考えられる場合にはより負担の軽い措置を義務内容にする」と主張する樋口亮介「注意義務の内容確定基準——比例原則に基づく義務内容の確定——」刑事法ジャーナル39号(2014年)50頁を参照。

せたにもかかわらず、社会がその残存リスクを許容しないのであれば、そのようなリスクを伴うIAの開発・利用は過失犯の実行行為として禁止されよう⁽¹⁶⁾。

(iii) 客観的帰属

グレス=ヴァイゲントは、背後者の行為と発生結果の帰属連関が否定される場合として、①一般的な生活リスクが実現したに過ぎない場合と②IAの動作が自律的な行為と認定される場合を提示する。しかし、①は、当該IAの開発・利用が「許されない危険」を創出するか否かを問うものであり、社会による当該リスクの受容の有無(上記(ii))の問題に還元される⁽¹⁷⁾。また、②については、予見可能性を検討した際にIAによる侵害結果を——その運用上想定できる種類の侵害結果の範囲内においては——予見できると解する限り、たとえIAが自律的に作動するものであったとしても、背後者の行為と発生結果の帰属連関(または相当因果関係)が否定されることはほとんどないだろう。

(iv) まとめ

予見可能性や客観的帰属の否定を通じてIAを利用する背後者の過失

(15) IAのリスクが極めて軽微であり、注意義務の履行によりその質・量を減少させるまでもなく、社会が当該リスクを受け入れるのであれば、IAの開発者・利用者に注意義務が課されることはなく、また、その使用に基づく侵害結果について刑事責任を負うことはない。

(16) ある行為の遂行に伴うリスク、注意義務の履行によるリスクの減少、減少したリスクとその行為のベネフィットの比較衡量のすべてを「実行行為論」の枠内で検討する以上の理解に対して、実行行為性(構成要件該当性)を検討する場面では注意義務の履行が実行行為性を失わせる程度にリスクを減少させるかを問い、注意義務の履行にもかかわらず残存したリスクとベネフィットの比較衡量は違法論で行うとの理解もあり得よう(こうした考え方を示唆するものとして佐伯仁志『刑法総論の考え方・楽しみ方』(有斐閣、2013年)309頁以下)。もっとも、ここでのベネフィットはIAの開発・利用がもたらす抽象的・一般的な利益なのであって、当該IAが侵害した個別具体的な侵害利益と比較衡量しうるのか、あるいは——比較衡量は可能であったとしても——侵害結果を正当化しうるのかは疑問である。

(17) 山中敬一『刑法総論〔第3版〕』(成文堂、2015年)400頁参照。

責任が否定される場合もありうるし、それらが過失犯の成否にとって重要な要素であることに争いはない。

しかし、IAのような先端技術を開発・利用する者の罪責を適切な範囲に限定しようと試みる場面においては、開発者・利用者により創出されるリスクが社会的に「許された危険」であるか否かが最も重要な問題になると思われる。予見可能なあらゆる侵害結果について開発者らに刑事責任を問うのでは、刑法が技術革新を停滞させ、社会全体の利益を損なうことになりかねない。しかし、社会生活の向上を目的とするならば如何なるリスクを伴う開発行為も法的に許されると解することもできない。それゆえ、一方で、開発者らは、開発段階においては——合理的なコストの範囲内で——当時の技術水準的に十分なテストを実施し、販売・利用段階においても製品の監視、場合によっては改良を行うことで、当該製品に伴うリスクを可能な限り減少させる必要があるが、他方で、開発者らが十分な注意義務を履行したにもかかわらずリスクがなおも残存する場合、社会が——当該製品が社会にもたらしうるベネフィットを理由に——それを許容する限り、当該リスクは社会全体に負担させるべきである。したがって、開発者らが過失犯として処罰されうるのは、①リスクの減少を内容とする注意義務を履行することなく製品を開発・利用する場合、あるいは②製品に伴うリスクが注意義務を履行したとしても社会的に許されるレベルには至らないにもかかわらず、当該製品の開発・利用を継続する場合に限られることになる。刑法は、以上のような形でIAを含む先端技術のリスクとベネフィットを比較衡量することで、「技術革新の促進」と「社会的安全の確保」という2つの要請を満たすべきであると思われる。