

2025年度

SC

小論文

3月12日(水)

人文社会科学部(法学科)

10 : 00 ~ 11 : 30

【後期日程】

注意事項

試験開始前

- 1 監督者の指示があるまで、問題冊子、解答用紙、下書き用紙に手を触れてはいけません。
- 2 監督者の指示に従って、全部の解答用紙(2枚)に受験番号を記入しなさい。

試験開始後

- 3 この問題冊子は、4ページあります。はじめに、問題冊子、解答用紙、下書き用紙(1枚(表裏))を確かめ、枚数の不足や、印刷の不鮮明なもの、ページの落丁・乱丁があった場合は、手をあげて監督者に申し出なさい。
- 4 解答は、すべて解答用紙に記入しなさい。(下書き用紙と間違わないよう十分注意してください。下書き用紙は採点対象となりません。)
- 5 文字数制限のある解答用紙の記入については、下記の点に留意すること。

- ・問1・問2の書き出しは、一マスあけないが、問3はあける。
- ・段落を改めるときは、最初の一マスをあける。
- ・句読点及び括弧等は、それぞれ一マス使う。行の末尾については文字と同じマスに含める。
- ・小さな文字「っ」「ゃ」「ゅ」「ょ」等は、それぞれ一マス使う。
- ・英数字は一マスに2文字入れてよい。

- 6 問題は、声を出して読んではいけません。
- 7 配点は、比率(%)で表示してあります。

試験終了後

- 8 問題冊子と下書き用紙は、必ず持ち帰りなさい。

次の問題文を読み、以下の設問(問1・問2・問3)に答えなさい。

[問題文]

AI(注1)をはじめとするデジタル技術に対しては、人間の能力を拡張する可能性に大きな期待が寄せられる反面で、想定外の動作をしたり、制御不能に陥ったりして、周囲に損害を発生させるという危険性が、常に問題とされてきた。特に、自動運転車や介護ロボットのように、人間と物理的に接触する可能性が高い技術の場合、このリスクは深刻なものになる。公道上で自動運転車の走行実験が行われているアメリカでは、すでに数十件を超える事故事例が報告されている。2018年3月には、アリゾナ州で実験走行中だったウーバー(注2)の自動運転車が、車道を横断しようとした人を無視してはね、死亡させるという人身事故も発生してしまった。

物理的な損害を伴わない場面でも、AIが判断を誤って被害を発生させるという問題は、放置できない。たとえば、医師が画像診断のAIを用いて医療を行ったにもかかわらず病気の兆候を見落とすケース、あるいは採用人事で志望書類をAIに読み込ませた結果、本来であれば採用されるべき人材が落とされてしまうケース、さらには、金融機関の導入したAIによる融資審査システムが、融資基準に合致しているはずの対象者を不適合と判断して融資を拒絶するケースなど、「AIの間違い」が原因となる問題事例は、いくつも想定することができる。

このような「間違い」は、あってはならないことであると同時に、万一発生した場合には、損害賠償責任が問題になる。ところが、モノの取引を基本にした現在の法の体系は、AIの特徴を十分に受け止めることができない。そこで、新しい考え方を生み出す必要性が強く感じられている。その一つとしてしばしば主張される解決策に、人間の判断を介在させるというものがある。AIに完全に依存してしまうのではなく、いわばそこにセーフティ・ボタンを付けておいて、ぎりぎりの場面では、人間がそれを押すことにすれば被害の発生を防げるのではないかという考え方である。

日本の裁判所による判決の中に、同じような考え方をとったものがある。AIの事件ではないが、コンピュータ・システムの不具合が大きな損害につながった事例である。

この事件は、証券会社が証券取引所に対して注文を出す際に、入力ミスをしたことが発端であった。問題の株式は、その日に上場されたばかりであり、1株67万円という気配値がついていた(買い注文の方が多かったため売買が成立しなかった。こういうとき、十分な売り注文があれば取引が成立しそうな価格が「気配値」として表示される)。証券会社の担当者は、この株式について、「61万円で1株の売り注文」を出そうとしたのであるが、うっかり間違えて、「1円で61万株」の売り注文を入力してしまった。1円の売り注文は明らかに不自然なので、画面上に、注意を促す警告サインが出たらしいが、それを無視して発注を確定させてしまった。67万円の気配値がついていた株式が1円で売りに出されたため、すぐに買い注文が殺到した。ここまでは、単なる人為的なミスで、システムの不具合の問題ではない。

間違いに気づいた証券会社は、この売り注文を取り消そうとしたが、証券取引所のプログラムにバグがあったため、取り消しが入力できなかった。他方、証券取引所の職員も異常に気づき、発注元の

証券会社に電話をかけて連絡を取ろうとしたがうまくつながらず、混乱しているうちに、システム上、売買が次々に成立していってしまった。結局、証券取引所の側では売買停止措置をとるという判断に至らず、誤った注文を出した証券会社が、自分で買い注文を入れて元の売り注文を解消させるという行動をとり、とりあえず、その場は收拾された。この間、合計8分あまりであった。

証券会社は、1株だけを売るはずが、膨大な数量の株式を売却したのとして決済しなければならなくなり、その分の損害を被ったとして、証券取引所に対し、損害賠償を請求した。もともとのミスをした張本人は証券会社側の担当者なのであるが、発注取り消しができていればこのような事態にはならなかったし、証券取引所が早い段階で売買停止措置をとっていれば、損失は小さかったはずだという主張である。この訴訟に対する判決の中で、東京地方裁判所は、最初の誤った売り注文から7分を経過した時点までには、証券取引所が売買停止措置をとることができたはずであると述べ、それをしなかった以上は、「人的な対応面を含めた全体としての市場システムの提供について、注意義務違反があった」と判断した。市場での取引自体は、プログラムにもとづいて執行されるシステムになっているが、そのコンピュータ・システムの運営は、証券取引所の職員という人的な組織が担っている。この事件で裁判所は、コンピュータ・システムの運用と、異常が認められたときは売買を停止するという人間の判断とを総合した全体が「市場システム」であり、その「システム」が全体として、損害の発生を防ぐように構築されていなければならないという考え方を採用したと言えよう(東京地裁平成21年12月4日判決。この事件は控訴され、控訴審の東京高裁ではこれ以外の論点によって結論が少し変わったが、ここでは立ち入らない)。

しかし、万一の場面で介入を求められる人間は、的確な判断にもとづく対応をとることができるであろうか。完全な自動運転に至る前の、レベル3の自動運転車(注3)でも、特に異常がなければ、人は運転操作をする必要がない。そのため、自動車の中で携帯電話やスマートフォンを使ってもよいことになった(令和元(2019)年改正による道路交通法71条の4の2)。とはいえ、その状態でも、何かの事態が発生したときは、突然のアラームで、自動運転装置から運転を引き継ぐように求められる。技術開発の現場では、この場合のアラームとして、たとえば、警告音を鳴らすよりも、シートなどを振動させる方が、より素早い反応を引き出せるといった実験も行われているが、どれほど素早く反応したとしても、引き継いだ瞬間にとるべき操作がとっさにわからず、まごまごしてしまうという可能性がないとは言えない。将来、さらに自動運転のレベルが進んで、運転操作をしないことが当たり前になると、何もしない状態に慣れきってしまい、突然に操作をしようとしても、パニックを起したり、かえって間違った操作をしてしまう危険性が高くなっていくであろう。

もちろん、乗っている人(運転者)が危険に気づいてもどうすることもできないということがないように、システムを止めて手動に戻す仕組み(いわゆるキル・スイッチ)はあった方がよい。AIシステムが広く普及しても、手動に戻すスイッチは残されるはずだと語る技術者は多い。しかし、そのことと、手動に戻す仕組みがあれば、法的な意味で安全性が確保されていると評価できることとは、まったく別の問題である。人間の介入が的確なタイミングで行われ、かつ、介入した結果が、AIに任せ

たまの状態で危険を抑える方向に働くという仕組みでなければ、安全性を確保したシステムとは言えない。このように考えると、人間の判断を介在させるとか、最終的に人間が操作を引き継ぐ余地を残すという方法で安全性を確保するというアプローチは、アイデア倒れに近いように思われる。

人身被害が生ずる事故のケースを離れて、一般的なAIの誤作動について考えると、問題の所在が、より明確になる。たとえば、医療診断や会計監査で膨大なデータを解析する際に、機械学習をしたAIを活用し、問題のあるデータを識別することが有効だと期待されている。この場合の機械学習では、AIが、既存のデータ（「教師データ」^(注4)）の中から、ある条件（たとえば病気の兆候を示す患部の写真）に合致するものと合致しないものを識別するパターンを学習する。その上で、新しいデータをAIに読み込ませ、学習した条件に合致するものを適切に選び出させる。医療の場合は、検査画像から病気の兆候を示すものをAIが判別し、会計監査では、財務諸表や会計帳簿のデータから粉飾を疑わせるものを判別する。これらの場合、人間が画像を見たり数値を読んだりするスピードと比較して、AIが画像データや数値データを読み込むスピードは、けた違いに早い。そこで、機械学習の結果、問題のありそうなデータの判別が一定以上の精度に到達すれば、医師や公認会計士の経験と勘に頼るよりもずっと効率的になると想定されているのである。この場合にも、AIが病気や粉飾を見落とすという間違いをしては困るので、人間が関与して、いわばダブルチェックを行うとよいのではないかとと言われると、一見、納得してしまいそうになる。

ところで、「間違い」の中にも、二種類の違ったものを区別しなければならない。本当は病気や不正があるにもかかわらず病気や不正はないと判断してしまうという間違い（第一種の過誤）と、本当は病気や不正ではないものを病気や不正と判断してしまうという間違い（第二種の過誤）である。AIの間違いをチェックするために人間が関与するべきだと言われる場合の「間違い」は、第一種の過誤の方であろう。病気の兆候が発見されれば医師は治療をしなければならず、公認会計士は粉飾の疑いが発見されればさらに調査を行うので、第二種の過誤は、通常の業務プロセスの中では問題にならないからである。

ところが、ここで第一種の過誤を防ぐために人間のダブルチェックを要求すると、せつかくAIを導入した意味が大きく損なわれてしまう。⁽²⁾というのは、AIのメリットは人間よりも速く、大量にデータを処理できる点にあるので、そのメリットを生かそうとすれば、人間に読めるような画像や文書に変換せず、データのまま処理するという方法をとることになるはずだからである。AIが詳しく見る必要がないとして捨てたデータを、その中に、実は見落としてはいけないデータがあったのではないかという理由から、人間に読めるような画像や文書に変換して、改めて人間がチェックするので、データ処理の効率はまったく上がらない。ここでも、人間が判断に関与することでAIの間違いを防ぐという考え方は、アイデア倒れに終わっている感がある。

AIの判断を人間がチェックするという考え方が現実的ではないとすれば、AIシステムの間違いや暴走に関する制度は、どのようにしていくことが理想であろうか。改めて、証券取引所の誤発注事件

の判決を読み直すと、「システム」という言葉がカギになっていることに気づく。売買の発注を行うコンピュータ・プログラムもシステム(発注システム)なのであるが、それを組み込んだ上位の概念として「市場システム」という言葉が使われ、その中に、証券取引所のスタッフによる売買停止措置などの人的な対応も含むとされているのである。そこでは、機械的なシステムを、どのような形で社会的なシステムの中に組み込んでいくかという問題が提起されていると言える。

システム設計の専門家は、「冗長性」という考え方をすることが多い。冗長性とは、システムの安全性を確保するために、安全装置を二重、三重に備えて、いわば余らせておく(「冗長」にしておく)ことで、一つの安全装置が働かなかった場合にも、他の装置が機能してリスクを回避できるという意味である。AIの利活用についても、特定のAIシステムが間違いを起こした場合に、それが損害に直結することがないように、別のシステムが作動して事故を防止するといった冗長性のある設計が要求されると考えられる。そこにいう「別のシステム」の中には、人間の関与という要素も含まれる。その意味で、人間の判断を介在させて安全を確保するという考え方は、間違っていない。ただし、人間の判断が介在するといっても、AIによる判断の限界をふまえて、その範囲外にあたる場合の判断フローを適切に設定するなど、AIの使われ方を考える責任という問題なのである。

(3)

(出典) 小塚荘一郎 『AIの時代と法』(岩波新書, 2019)45-54頁。

なお、出題にあたり、原文にあった小見出しを省き、原文にない(注1)、(注2)、(注3)、(注4)を追加し、漢数字をアラビア数字に置き換え、文章を一部改めた。

(注1) AI Artificial Intelligence の略。人工知能のこと。

(注2) ウーバー 配車サービスのこと。タクシー運転手でない者でも、空いた時間を利用して乗客を目的地まで送り届け、対価を得ることができる点に特徴がある。

(注3) レベル3の自動運転車 システムがすべての動的運転タスクを限定領域において実行する運転のこと。システムの作動継続が困難な場合は、運転者が介入要求等に適切に応答する必要がある。

(注4) 教師データ AIに学習させる基盤となるデータのこと。問題と正解がセットになっており、AIにこのデータを学習させることで、正しい予測の精度を向上させる。

[設問]

問1 下線部(1)の文章について、その内容を150文字以内で説明しなさい。(配点20%)

問2 下線部(2)の文章について、「AIを導入した意味が大きく損なわれてしまう」のはなぜか。その理由を150文字以内で説明しなさい。(配点20%)

問3 問題文の内容を踏まえて、下線部(3)について自分自身の考えを600字以内でまとめなさい。(配点60%)