

奨励金No.1526

人工知能を使った裁判支援システムに対する 一般市民の評価—「AI 裁判官」は 受け入れられるのか？

綿村 英一郎

大阪大学大学院・人間科学研究科 准教授

The Public's Evaluation of Trial Support Systems Using Artificial Intelligence — Will “AI Judges” be Accepted?

Eiichiro Watamura,

Graduate School of Human Sciences, Osaka University, Associate professor



本プロジェクトでは、裁判においてAIを使用することが人々にどの程度支持されるのか、また、その支持を規定する要因は何かという2つの問題について検討した。2022年9月に行ったWeb調査の結果、その支持は1～5ポイント中2.64と全体的に低く、裁判におけるAIには否定的であることが示された。また、AIが非人間的で不透明であるとみなされるほど支持されなくなり、逆に公平だとみなされるほど支持されていた。今後は、この調査のフォローアップと残りの実験計画を進めていく予定である。

This project examined the extent to which people support using AI in criminal courts and the factors related to this support. The web survey conducted in September 2022 showed a low mean score of support (2.64 out of 1–5 points), indicating that using AI in criminal courts is considered negative. It also showed that the more “impersonal” and “opaque” the AI is perceived, the less support it receives. Conversely, the more “fair” it is perceived, the more support it receives. We intend to follow up on this survey and plan the rest of the experiments.

1. 研究内容

1. 目的

本プロジェクトでは、人工知能による裁判支援システムに対する人々の反応について検討した。今やAI (Artificial Intelligence) はあらゆる分野で用いられる必要不可欠な技術となりつつある。裁判も例外ではない。実際、アメリカの一部の州ではCOMPASと呼ばれる再犯リスク予測システムが導入されており、中国でも判例データから量刑判断を補助するシステムが導入されている。日本でも、事案を入力することで法的推論を行う裁判支援システムの開発が行われている (二木、2022)。認知バイアスや感情の影響を受けやすい人

間と比べ、AIの判断は一貫して客観的であり、たしかな優位性がある。しかし、医療診断、経営支援、自動運転といった場面でのAI化は歓迎ムードであるのとは対照的に、裁判支援を担うAIの導入に対しては違和感を覚える者も多い。太田 (2020) が一般市民に行った調査によれば、AIによる裁判支援システムには期待感もあるが不安感も高かった。それでは、裁判においてAIを使用することはどの程度支持されるのだろうか。また、その支持を規定する要因は何であろうか。

2. 調査的研究

2.1. 概要

2022年9月、インターネット調査会社に登録された一般市民を対象とした意識調査を行った（特定の年・性別にサンプルが偏らないように、人口推計に沿って性別と年齢（5歳区切り）を割り当てた）。1,000名を募り、基準を満たさない一部のサンプルを除いた783名（女性423名、男性360名、平均年齢52.3歳）の回答を分析した。

2.2. 質問項目

裁判におけるAI使用への「支持」を尋ねた。具体的には、「現在、たとえば犯罪をおかした人を裁く裁判では、人間の裁判官が判断を下しています。しかし近年、このような裁判を人間ではなく、AI（人工知能）によって行う国が登場しており、日本でもその実用性の検討がなされています。あなたはAIの裁判官によって人を裁くことについてどのように考えますか」というリード文を提示した。その後続けて「裁判にAIを導入するのは望ましいことだ」などの4項目に対して「1：全くそう思わない」から「5：非常にそう思う」までの5件法で回答を求めた。裁判におけるAI使用への「イメージ」を尋ねた。「AIを使えば、人間の裁判官よりも公平な判断をすることができる」といった公正さに関する5項目、「AIはあまりに複雑なので、裁判で使うと、どのような理由で判決を下しているのか分からなくなる危険がある」といった不透明さに関する5項目、「AIは、人間の裁判官のように『人情味』のある判決を下すことはできない」といった非人間性に関する5項目であった。これら15項目に対して、5件法で回答してもらった。さらに、「AIが感情を持てば不安だ」といったAI人工物否定的態度（古賀、2022）、AIに関する知識や経験を統制変数として尋ねた。

2.3. 結果と考察

探索的因子分析を行い、各項目群の因子構造を

調べた。その結果、裁判におけるAI使用への「支持」は1因子、裁判におけるAI使用への「イメージ」は3因子であり、内訳は、AIに対して肯定的な「公平」因子、否定的な「非人間性」因子と「不透明」因子の2つであった。人工物否定的態度はポジネガの2因子であった。いずれの因子についても十分な内的一貫性が確かめられたため、それぞれを平均値化したうえで以降の分析を進めた。重回帰分析の結果、裁判におけるAI使用への「支持」に対して、非人間性、不透明が負の関連を示し、公平イメージは正の関連を示した。つまり、AIが非人間的で不透明であるとみなされるほど裁判でAIを使うことが支持されなくなり、逆に公平だとみなされるほど支持された。この結果は、直感的にもそれらしいと考えられる。「支持」とその他の一部の変数にも関連が見られた。この調査で興味深いのは、裁判におけるAI使用への「支持」がそもそも低いという点であった。分析対象となったサンプルのスコアの平均値、95%信頼区間は2.64 [2.59, 2.69]であった。理論的中点は3点であったから、調査時点の市民の意識としては裁判におけるAIには否定的であったといえるだろう。医療や自動運転分野では必ずしももうならないことを鑑みると、人間の安全や生命以上に倫理や道徳的な判断は重要であり、人間の専売特許であるべきだとみなされている可能性が示唆される。

3. 今後の課題

このプロジェクトにはいくつかの課題が残っている。まずはこの調査的研究の補強である。特に、サンプルがインターネット・ユーザーに限られていた点には問題がある。パソコンであれスマートフォンであれ、情報通信機器が使いこなしている層であれば、全く使わない層に比べ機械に対する違和感が低いと想定されるためである。ただし裏を返せば、そういう層でさえ裁判におけるAI使用はあまり支持されてないといえる。さらに、最近のAIについてどのような知識や先入観をもってい

るのかという切り口も重要である。たとえば、ChatGPTについて何らかの脅威を抱いていれば、裁判場面での応用にも慎重になるだろう。最後の課題は実際のAI裁判官を提示したうえでの実験的検証である。「AIの裁判官によって人を裁くことについてどのように考えますか？」と尋ねられても、人々が想像するAIは人それぞれだろう。映画「ターミネーター」のような残虐性のあるAIを想像する人もいれば、「ベイマックス」のようなコンパニオンロボットを想像する人もいるだろう。未来の法廷で導入されるAIが実際にどのような形態になるかをふまえたうえでの検証が必要だろう。

2. 発表（研究成果の発表）

向井智哉・湯山祥・綿村英一郎、裁判におけるAI使用への支持の程度と規定要因、情報知識学会誌、33巻1号、3-19、(2023年)