法とAIの心理学
――AI裁判官は受け入れられるのか?

大阪大学大学院基礎工学研究科／経営企画オフィス 助教
井奥智大

いおく・ともひろ　博士（人間科学）。専門は社会心理学と人間科学。2025年より現職。筆頭論文にTrade-offs in AI assistant choice: Do consumers prioritize transparency and sustainability over AI assistant performance? *Big Data & Society, 11*(4), 2024など。

近年，人工知能（AI）の進展に伴い，その応用範囲は司法分野にも拡大しつつある。特に，被告人に対して判決を下す「AI裁判官」は，技術的に実現できるかどうか，実現できたとしてそれが正しいことなのかという技術的・倫理的問題から関心を集めている。しかし，こうした革新的技術の社会的受容や，その根底にある心理的メカニズムについては，依然として実証的知見が限られている。これまでの研究はAIアシスタントとしての活用に焦点を当てており，AIが裁判官として自ら判断する場合に人々がどのように感じ，評価するかは十分にわかっていない。

AI裁判官は人間を超えられるか

AI裁判官は，人間の判断にしばしば伴う「ノイズ」や認知バイアスを排除し，一貫性のある判断を下す潜在的能力を有している。また，膨大な法的文書や判例情報を瞬時に処理・参照できるため，情報処理能力の面でも人間を上回る可能性が指摘されている。しかし，こうした優位性にもかかわらず，AIによる判断過程が外部から理解しづらい「ブラックボックス」問題などの懸念が存在する。実際に，筆者が日本で行った研究でもAIの性能の高さだけでなく，「どうやって判断しているのか」という透明性が，AIを受け入れるかどうかに大きく関わることが示されている[1]。

他方，自動運転など他の分野の研究では，AIが利用者に共感を示すことで，そのAIへの信頼や受け入れやすさが高まることが示唆されている。これは裁判という場面にも当てはまるかもしれない。裁判では，客観性が何よりも重視されるが，同時に，当事者への共感も評価に関わる場合がある。もしAI裁判官が人間のような共感を持ちながら高い客観性を保つことができれば，人間とAIそれぞれの強みを合わせた理想的な裁判官を実現できる可能性がある。

映像実験から得られた知見

AI裁判官への反応を調べるため，筆者らは刑事裁判を再現した約5分間の短い映像を4種類作成し，オンライン実験を行った[2]。映像は「裁判官の種類」（人間／AI）と「共感の有無」（共感あり／共感なし）を組み合わせた2×2の条件で構成された。実際の裁判場面を模した映像を用いることで，文章だけでは捉えにくい裁判官の表情や口調といった非言語的な情報も含めて提示でき，より現実に近い形で参加者の認識や感情を測定できる。

研究の結果，共感操作の有無によって，参加者の反応に一貫した差異がみられた。具体的には，共感あり条件では，裁判官（人間，AIを問わず）に対する信頼が高まり，判決評価やAI裁判官の受容の高さに関連していた。一方で，同じ共感的発言であっても，人間裁判官の方がAI裁判官よりも高い共感性を持つと評価される傾向が示された。

受け入れられるAI裁判官とは

将来，AI裁判官を司法制度に導入するには，AIの客観性だけでなく，人間らしい共感といった心理的要素を高めることが，社会に受け入れられるための重要なポイントになる。そのためには，AI裁判官における「共感」をどう定義し，どこまで倫理的に認めるのかを，法律・心理・倫理などの分野の専門家が協力して話し合う必要がある。また，著者らが行った別の研究では[3]，刑事裁判において裁判官とAIが異なる量刑判断を提示しても，裁判員はどちらか一方に偏ることなく，情状酌量の有無をより重視して判断していることが示されている。映像実験の知見と併せて考えると，これはAIと人間裁判官が対等に意思決定の参考になり得る可能性を示唆すると同時に，AIに共感を備えられれば，客観性と人間らしさを兼ね備えた「ハイブリッド型裁判官」の実現に近づくことを意味する。このような知見はAI裁判官の設計や社会実装に向けた議論の基盤となるだろう。

1 Ioku, T. et al. (2024) *Big Data Soc, 11*(4). https://doi.org/10.1177/20539517241290217　**2** Watamura, E. et al. (2024) *Int J Hum-Comput Interact, 40*, 5192-5201.　**3** Watamura, E. et al. (2025) *PLoS One, 20*, e0318486.