

検証結果の要旨

1 検証の方法等

東京地裁及び大阪地裁において民事訴訟を担当する中堅裁判官 6 人が参加し、模擬記録を用いて、争点整理の補助としての生成 AI の活用について、検証を行った。

具体的には、各参加者から事前に争点整理及びそれに付随する範囲の場面の事務における生成 AI 利活用のアイデア（ユースケース）の提出を受けた上で、1 月 20 日及び 21 日に第 1 回ワークショップを、2 月 17 日に第 2 回ワークショップをそれぞれ開催し、外部業者の協力を得ながら、3 つの生成 AI サービス（ChatGPT 5.2 Thinking、Gemini 3 Pro（思考モード）、NotebookLM Pro）を用いて、争点整理の補助として利活用ができるかどうか、利活用に当たりリスクはあるかなどについて検証を行った。

2 AI 活用が特に有効と考えられる場面

- ① 主張整理表や時系列表等のたたき台作成
- ② 大量の主張書面や証拠の要約等
- ③ 引継ぎ時等における全体像の迅速把握
- ④ 文字起こし、要約

3 課題・リスク等

(1) 出力品質に関する課題

① 抜け漏れの発生

重要な事実や証拠の拾い漏れが一定程度存在し、AI の出力のみに依拠した場合、重大な見落としにつながりかねないとの懸念が共有された。（検証参加者からは「必ずしも重要な間接事実を全て拾えているわけではないと思われる上、間接事実を裏付ける証拠も全て拾えているわけではない」「網羅性に問題が大きいと、記録検討の手間はさほど軽減されない」との意見が出された。）

② 出力の揺れ（再現性の低さ）

同一のプロンプト・同一の入力データに対しても、実行のたびに出力結果が異なるという問題が多く、検証参加者から指摘された。

③ 期日ごとに情報が更新されていくことへの対応の弱さ

新たな書面を追加した場合に、既存の整理結果が崩れる、あるいは精度が大幅に低下するという問題が報告された。検証参加者からは「各期日で提出された書面を追加して、主張整理を更新していくと、前の期日

までの整理に引っ張られるのか、主張整理の正確性が大幅に低下した」
「訴状と答弁書のみを対象とした際には、ある程度正確に争点を把握できていたが、その後、準備書面を追加すると、争点の把握として不適切な部分が出てきた」との具体的な報告がなされた。

④ 法的理解の不十分さ

規範的要件の位置づけの誤り、特殊な抗弁構造の理解困難、攻撃防御方法の体系的整理における不正確さ等が複数のユースケースで報告された。この課題は、主張分析・争点整理支援のユースケース全体に通底するものであり、生成A Iの法的推論能力の限界を示唆している。検証参加者からも「要件事実の理解が不明確／不十分」「要件事実を正しく理解させるために誘導や指示が必要」との意見が共通して出された。

(2) 実務上のリスク

ア 誤導・バイアスの危険

「注意すれば使用可能」との評価が大勢であったが、一部のユースケースでは具体的な誤導リスクが指摘されている。生成A Iの出力が「もっともらしい誤り」を含む場合、裁判官の先入観を助長するリスクがあり、常に批判的な検証が不可欠である。検証参加者からも上記誤導リスクからすると「証拠との不整合指摘に利用するのはリスクが高い」との見方も示された。

イ 過信のリスク（検証姿勢の弛緩）

A Iの出力がきれいに整理された形式で提示されるため、内容の正確性に対する信頼が過度に高まる危険性がある。検証参加者からも「便利なものであればこそ、それしか見なくなってしまうおそれがある」との懸念が表明された。

ウ 倫理面のリスク

A Iの整理結果を参照する過程で、裁判官自身の主体的な検討が不十分なまま、A Iの出力に沿った方向で判断の枠組みが形成されるおそれがある。

エ 運用上の課題（スキル格差・能力維持）

A Iの利用スキルの個人差により、裁判官間で生成A I活用の効果に差が生じ得るとの指摘や、定型的な整理作業をA Iに委ねることで、裁判官自身の記録検討能力や法的分析能力が長期的に低下するリスクが指摘された。検証参加者からも「対外的な受け止めや裁判官の能力の低下の問題等も含めて慎重さも必要」との意見が出されている。

(3) 情報セキュリティ面のリスク

本検証は模擬データを用いて実施されたため、実際の訴訟記録に含まれ

る要機密情報の取扱いに関するリスクは直接的には検証されていない。もっとも、今後の本格検証及び実運用に向けては、情報セキュリティ上の課題について、慎重な検討が必要である。

4 対応策等

① 裁判官が出力結果の確認を行い、最終判断と責任は裁判官が負うこと

A I の特性として抜け漏れ等が生じることから、生成A I の出力は、裁判官による検討を経ずにそのまま依拠するものではなく、裁判官が出力の正確性を確認し、必要な補正を行うことを前提として利用するものであり、出力に対する裁判官の批判的な検証が不可欠である（ヒューマン・イン・ザ・ループ）。検証参加者からも「人による記録検討に基づく検証が必須」との認識が共有された。

② 必要に応じて当事者とも認識共有を行うこと

A I の特性として抜け漏れ等が生じることから、出力結果と当事者の主張・立証との間に齟齬がないかどうかにつき、必要に応じて当事者と認識共有をすることも重要と思われる。

③ 法的枠組み（訴訟物、攻撃防御方法の構造、争点等）の設定は裁判官が行うこと

現状では生成A I は要件事実論に基づく自律的な法的整理が困難である上、法的枠組みの整理は裁判官が行うべき事柄であると考えられるため、裁判官が法的枠組みを指定した上で、A I にその枠内での整理作業を委ねるという役割分担が必要となる場合がある。

5 本格検証に向けた課題

(1) 実環境に近いデータを用いた検証

本検証は模擬データ（学修用記録）を用いたものであり、比較的シンプルかつ整理された記録が対象であった。複雑困難な事件や当事者の主張が整理されていない事件において実用に耐え得るかどうかは不明であり、確定記録等の実際の訴訟記録を用いた検証が不可欠である。

(2) 情報セキュリティ等の整理

実際の訴訟記録を用いた検証及び運用に向けては、適切な利用環境の整理が不可欠である。訴訟記録には当事者のプライバシーに関わる情報や営業秘密等が含まれ得ることから、十分なセキュリティ措置が確保された環境での実施が前提となり、適切な利用環境を設計する必要がある。

以上