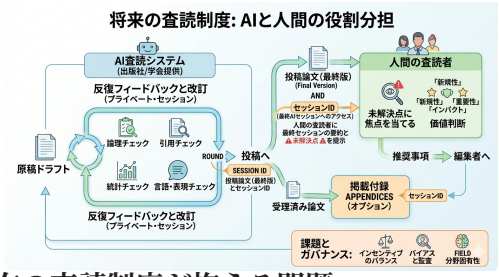
AI 時代の論文査読はどう変わるのか

著者：関 勝寿　公開日：2026 年 6 月 18 日　ソース：https://sekika.github.io/2026/06/18/AI-review/

引用：関勝寿 (2026). AI時代の論文査読はどう変わるのか . Zenodo. https://doi.org/10.5281/zenodo.20763535

近年、AI を論文査読に利用することの是非が盛んに議論されている。しかし、その多くは現在の AI の性能を前提としている。確かに現時点では AI による査読にも限界があるが、制度設計を考える上で重要なのは現在ではなく将来である。AI の能力は今後も向上していく一方で、学術出版の制度は容易には変わらない。そこでこの記事では、AI が査読プロセスに本格的に組み込まれた場合、人間査読者と AI がどのように役割分担することになるのかを考察する。



1. 現在の査読制度が抱える問題

現在の査読制度では、人間査読者が非常に多くの役割を担っている。査読者は論文を読んで、論理的な誤りや説明不足を探し、引用文献の妥当性を確認し、統計解析や実験設計の問題点を検討する。同時に、その研究が本当に新しいのか、どの程度重要なのか、掲載に値するのかも判断しなければならない。

しかし近年は、査読負担そのものも増加している。大規模言語モデルの普及によって論文執筆のコストが低下し、研究者は以前よりも容易に論文を書けるようになった。AI が研究活動を支援するようになれば、この傾向は今後さらに強まるだろう。その結果、投稿される論文数は増加する一方で、それを査読する人間の数が同じ割合で増えるとは考えにくい。

もし AI によって論文を書くことが容易になるのであれば、論文を読む側、すなわち査読の側にも AI を導入しなければ、学術出版システム全体のバランスが崩れてしまうかもしれない。

2.AI は査読プロセスの前段を担う

そのような背景から、将来的には、論文投稿の前段階に AI 査読が組み込まれる可能性がある。

ただし、それは著者が自由に選んだ AI を使うという話ではない。論文の機密性や査読の公平性を考えれば、出版社や学会が公式の AI 査読システムを提供する形になるだろう。

著者は論文を書き終えたら、その出版社が提供する AI 査読システムに原稿を提出する。AI は論文を読み、論理の飛躍、説明不足、引用漏れ、統計処理に関する疑問点などを指摘する。著者はその指摘に応じて修正を行い、再び AI に確認させる。このやり取りは何度繰り返してもよい。

現在でも著者は共同研究者や同僚からコメントを受けながら原稿を改善している。AI との対話は、それをより体系的かつ継続的に行う仕組みと考えることができる。なお、この草稿段階での往復は、いわば自由な相談であり、その内容自体が提出物の一部になるわけではない。

3. 提出するのはログではなくセッション ID

草稿段階での AI とのやり取りは数十回、場合によっては数百回に及ぶかもしれない。そのすべてを人間査読者が読むのであれば、かえって負担が増えてしまうし、そもそも試行錯誤の過程を査読者に見せる必要はない。

そこで、提出物に含めるのは草稿段階のやり取りではなく、著者が「完成した」と判断した段階で実行する、もう一段階の AI 査読である。この段階で AI に提出される原稿は、その時点での完成版そのものである。したがって、AI がここで新たな問題を指摘した場合、著者に残された選択肢は一つしかない。それは、その指摘が妥当でない理由を説明することである。もし著者が AI の指摘を妥当だと考えて原稿に手を加えるなら、それは同一セッション内の「修正」ではなく、別バージョンの完成版を新たに作ることを意味する。その場合、著者は新しい完成版に対して、改めて最終確認の AI 査読を実行しなければならない。著者は、AI からの指摘がすべて説明によって解消できると判断できる段階に至るまで、この最終確認を繰り返すことになる。そして、それ以上修正も説明も必要ないと判断した時点で投稿を行う。

出版社のシステムには、この一連の試行のうち、実際に投稿された原稿に対応する最後の最終確認セッションだけが、論文と紐付けて記録される。それ以前の試行や草稿段階での相談の履歴は、提出物には含まれない。つまり、投稿時に提出されるのは長大なログではなく、この最後のセッション一回分を指すセッション ID のみである。

4. 人間査読者は未解決の論点を確認する

人間査読者はまず論文を読む。その後、必要に応じて、その提出段階の AI 査読セッションを参照する。特に重要なのは、AI と著者の間で最後まで見解が一致しなかった論点である。

AI は問題があると考えているが、著者は修正ではなく説明によってそれに応じた。そのような箇所は、人間査読者が重点的に確認すべき候補となる。ただし、この仕組みが著者の行動にどのような影響を与えるかについては、慎重に考える必要がある (この点は記事の最後で改めて取り上げる)。一方で、AI が指摘し著者が修正済みの問題については、査読者は詳細を追う必要がないかもしれない。

この仕組みによって、人間査読者は無数の可能性の中から問題点を探し出す作業ではなく、すでに整理された論点を評価する作業に集中できるようになる。

5. 査読者の役割は「誤り探し」から「価値判断」へ

このような仕組みが定着した場合、人間査読者の役割そのものが変わるだろう。

論理的な矛盾や説明不足の発見、関連文献の探索、統計処理の確認といった技術的な作業の多くは AI が担当するようになる。一方で、人間査読者はその研究が本当に新しいのか、その分野にどの程度の影響を与えるのか、掲載する価値があるのかといった判断により多くの時間を使うことになる。

もちろん、AI が見逃した問題を人間が発見することは今後もあるだろう。しかし、人間査読者の主な役割は次第に「誤りを探すこと」から「研究の価値を評価すること」へと移っていくのではないかと思う。

6. セッションログを論文の付録として公開する

提出段階の最終セッションは、出版社のシステム内に記録されるだけでなく、論文が採択された際には、その記録を付録として公開してもよいのではないか。

人間査読者によるオープン査読の議論では、査読者の身元や評価が公になることへの抵抗が、長年最大の障害となってきた。査読者は率直な批判によって著者から報復的な評価を受けることを恐れ、匿名性を求める。しかし、AI の査読セッションには、保護すべき「査読者のキャリア」が存在しない。オープン査読が抱えてきた最大の反対理由は、この設計には当てはまらない。

さらに、最終セッションでは著者に「修正」ではなく「説明」しか選択肢がないという設計と組み合わせると、もう一つの効果が期待できる。著者の説明が、編集者や査読者という限られた読者だけでなく、論文を読むすべての専門家の目に触れることになるからである。一人の査読者の判断だけに委ねるよりも、説明の妥当性に対する検証が分野全体に分散される。これは、一部の学術誌が査読コメントと著者の返答を公開している運用とも自然に接続する発想である。

ただし、公開を著者の自由選択 (オプトイン) とすれば、「公開していない」という事実そのものが何かを隠しているという疑念を生みかねず、結果的に事実上の強制力を持ってしまう。これを避けるなら、採択された論文については原則として公開する運用のほうが現実的だろう。一方、不採択になった場合や他の出版社に再投稿する場合にこのセッションをどう扱うかは、まだ十分に検討されていない。

そして、この公開という仕組みは、著者のインセンティブに関わる、より根本的な問題とも結びついている。それについては、次節で改めて論じたい。

ただし、ここまで述べてきた仕組みを実際に機能させるためには、まだ解決されていない課題が残っている。最後に、その主要な論点を挙げておきたい。

7. 今後議論が必要な論点

7.1. 著者のインセンティブと「迎合」のリスク

この設計では、提出段階の最終セッション内で AI と著者の意見が一致しなかった箇所が、人間査読者の重点確認対象になる。しかし、不一致そのものが査読者の注意を引くというルールは、著者にとって本音の反論そのものをリスクにしてしまう。確信があっても AI の指摘を表面的に受け入れて修正したほうが安全だと著者が考えるようになれば、科学的な議論の質は却って下がりかねない。

セッションログを論文の付録として公開する案は、この問題を解消するわけではなく、むしろ別の形で強める可能性がある。説明が公開されるとなれば、著者はその説明を「査読者一人を納得させるための言葉」ではなく「分野全体の目に晒される永続的な記録」として書くことになる。これは説明の質を高める方向にも働き得るが、同時に、AI との不一致を記録に残すこと自体の心理的なコストを増す方向にも働く。結果として「反論する」よりも「とりあえず修正してしまう」方向への圧力が、これまで以上に強くなるおそれがある。

これは現在の人間同士の査読にも形を変えて存在する問題だが、AI との対話が機械的に記録され、かつ最終的に公開される可能性がある仕組みでは、その圧力がより強く、より一律に働く可能性がある。不一致を「隠れたペナルティ」にしない運用上の工夫が必要になるが、これは査読文化や評価者教育に関わる問題であり、技術的に即座に解決できるものではない。

7.2. 単一 / 複数 AI システムによるガバナンスの問題

出版社や学会が公式の AI 査読システムを提供するという設計は、機密性と公平性の観点からは妥当に思える。しかしこれは新たな権力集中の問題を生む。一つの出版社が単一の AI ベンダーに依存すれば、その AI の系統的なバイアスや盲点が、その出版社が出すすべての論文の査読に影響する。逆に出版社ごとに異なるベンダーを採用すれば、分野や媒体によって査読基準が暗黙的に分裂しかねない。

どちらに進むにせよ、AI ベンダーの選定基準や評価の透明性、ベンダー間の整合性確保といった、技術よりもガバナンスに属する問題が前面に出てくる。これは一研究者や一出版社の判断では解決できず、学術出版界全体での合意形成を要する論点である。

7.3.AI 査読システム自体の評価と監査

公式の AI 査読システムを導入する場合、その AI 自体を誰がどのように評価するのかが問題になる。AI が著者の修正方針や人間査読者の注意配分に影響する以上、それは単なる補助ツールではなく、学術出版の制度的インフラの一部になる。

したがって、AI がどの種類の問題を発見しやすく、どの種類の問題を見落としやすいのかを継続的に評価する必要がある。また、特定の研究手法、理論的立場、文体、言語、地域に対する系統的な偏りがないかも検証しなければならない。さらに、モデル更新によって査読基準が暗黙に変化する可能性もあるため、バージョン管理、更新時の影響評価、第三者監査といった仕組みが必要になるだろう。

7.4. 分野ごとの差異

AI 査読の有用性やリスクは分野によって異なる。統計解析、実験設計、データ処理、報告ガイドラインへの適合性など、比較的形式化しやすい確認項目が多い分野では、AI 査読は有効に機能しやすいかもしれない。

一方で、理論的独創性、概念の再解釈、歴史的文脈、文献読解の細部などが重要な分野では、AI が既存の標準的な理解に寄りすぎる可能性がある。その場合、AI 査読は新しい議論を支援するよりも、既存の枠組みへの適合を促す方向に働くかもしれない。したがって、AI 査読を全分野に一律に導入するのではなく、分野ごとの査読文化や論文の性質に応じた運用が必要である。

7.5. 費用負担とアクセスの公平性

公式の AI 査読システムには、計算資源、保守、セキュリティ、監査などの費用がかかる。その費用を出版社、学会、著者、研究機関、助成機関の誰が負担するのかは重要な論点である。

もし利用料が投稿料や掲載料に上乗せされれば、研究資金の少ない著者にとって新たな障壁になる可能性がある。また、大手出版社や有力学会だけが高性能な AI 査読システムを整備できるようになれば、小規模ジャーナルとの格差も広がりかねない。AI 査読を制度化するなら、費用免除、共同インフラ、非営利基盤など、アクセスの公平性を確保する仕組みもあわせて議論する必要がある。