AI-Assisted Manuscript Review: Opportunities, Challenges, and Reviewer Support Framework

Dayan Adionel Guimarães

***Abstract*—This article explores the potential integration of AI tools into the manuscript peer review process. It presents a discussion on the pros and cons of AI-based and AI-assisted reviewing, a structured list of inputs required for AI-driven pre-evaluation, and prompt templates that can be used with AI systems such as ChatGPT. The aim is to support a hybrid review framework where AI acts as a preliminary evaluator, guiding subsequent human reviewers toward a more efficient and balanced assessment.**

***Index Terms*—Artificial Intelligence, Methodology, Peer review.**

I. Introduction

The peer review process is widely regarded as a cornerstone of academic publishing, ensuring that submitted manuscripts meet rigorous standards of quality, relevance, and scientific integrity. By facilitating critical evaluation by domain experts, peer review serves as both a selection mechanism and a source of constructive feedback for authors. However, in practice, the process is increasingly strained by growing submission volumes, limited reviewer availability, and lengthy review cycles. Editorial teams must frequently remind reviewers to meet deadlines, manage unresponsive referees, and reconcile inconsistent evaluations [1].

These challenges have led to growing interest in the use of Artificial Intelligence (AI) to contour parts of the editorial workflow. AI tools are already embedded in the publishing ecosystem for tasks such as plagiarism detection, grammar correction, and reference formatting. More recently, proposals have emerged to extend AI's role into more substantive stages of manuscript evaluation, including reviewer recommendation, quality assessment, and content screening [2]. While still in early adoption phases, these developments reflect a broader trend toward complementing human judgment with automated systems.

The integration of AI into peer review raises both opportunities and concerns. On one hand, AI systems, particularly those based on natural language processing, can rapidly extract structural features, assess completeness, and provide preliminary evaluations aligned with journal criteria. These capabilities offer potential gains in consistency, speed, and scalability, especially at the triage stage. On the other hand, there are significant limitations, as AI lacks true understanding of research novelty, domain-specific nuances, and ethical implications. Furthermore, algorithmic decision-making may reproduce or amplify hidden biases present in training data [3].

D. A. Guimarães is with the National Institute of Telecommunications (Inatel), Santa Rita do Sapucaí, MG, Brazil, e-mail: dayan@inatel.br. ORCID 0000-0002-1304-792X.

Rather than advocating for full automation of the manuscript review process, this article proposes a hybrid approach in which AI acts as a first-step evaluator to support subsequent human reviewers. Such a model aims to reduce reviewer burden while preserving the nuanced judgment required for high-quality scholarly assessment. The article aims to provide a structured framework for using AI in manuscript review, including its potential advantages and limitations, guidelines for input preparation, and prompt design for human-in-the-loop collaboration.

The scope of the discussions in this article are limited to the use of AI as a merit-based and technical tool, emphasizing that fairness and social concerns should be addressed by human reviewers or editorial boards, not by automated systems.

A. Related Work

Peer review has long been scrutinized for its susceptibility to bias and inconsistency. Lee et al. [4] provided a foundational analysis of how subjective judgments, institutional prestige, and reviewer fatigue can distort evaluative outcomes, calling for methodological rigor in understanding bias mechanisms. To complement this concern, Brown and Heathers [5] introduced the granularity-related inconsistency of means (GRIM) test, a simple but effective method for detecting mathematically inconsistent means in reported results, revealing the prevalence of unnoticed statistical anomalies. Although targeted to psychology literature, this test is applicable to other areas as well.

As interest in computational support grew, Mrowiński et al. [6] explored the use of evolutionary computation to assist journal editors in reviewer selection, proposing an AI-based system that optimizes reviewer-paper matching while minimizing cognitive and systemic bias. Heaven [2] later reported on early industry experimentation with AI-generated reviews, documenting both enthusiasm for improved efficiency and concern about the opacity and interpretability of AI evaluations.

Efforts to formalize AI transparency and accountability were advanced by Mitchell et al. [7], who proposed 'model cards' as a standardized documentation framework to disclose the assumptions, limitations, and performance metrics of machine learning models used in high-stakes contexts. Meanwhile, Horbach and Halffman [1] analyzed the reluctance of editorial boards to implement peer review reforms, despite recognizing systemic challenges such as reviewer overload and process inertia.

Recent contributions have increasingly framed peer review not only as a technical process but as a civic responsibility within academia. Squazzoni [3] emphasized that peer review should be institutionalized as part of the academic 'social contract', deserving formal recognition and reward. In parallel, Doskaliuk et al. [8] provided a field-specific account of AI's integration into biomedical peer review, highlighting the trade-off between increased efficiency and the risk of eroding human-centered oversight.

The reviewer guidelines published by *Nature Communications* [9] emphasize ethical responsibilities such as confidentiality, fairness, and professional accountability. While these principles are essential to the integrity of peer review, they involve normative reasoning and situational judgment that current AI systems are not equipped to replicate.

Mrowiński et al. [10] investigates how evolutionary algorithms, specifically cartesian genetic programming (CGP), could improve editorial decision-making and efficiency in the peer review process. The study empirically demonstrated that using CGP effectively optimized reviewer assignment and significantly reduced the time required for the peer review process, showing up to a 30% improvement without increasing the reviewer workload. While [10] contributes specifically with empirical results validating algorithmic efficiency, the present work contributes through an analytical and critical framework addressing AI's limitations, ethical issues, and epistemological concerns, not just efficiency metrics.

Checco et al. [11] examines the benefits, risks, and limitations associated with employing AI tools specifically for manuscript screening processes in scholarly publishing. The work critically analyzed various AI applications, such as plagiarism detection, preliminary manuscript quality checks, and filtering tasks, highlighting potential efficiency improvements alongside challenges, including algorithmic bias, reliability issues, and ethical considerations. On the other hand, the present work expands beyond screening, delving into AI's potential role in evaluating nuanced and complex tasks, including assessing scientific novelty, methodological validity, and significance. It incorporates a more comprehensive and nuanced discussion of epistemological implications, ethical considerations, and potential risks associated with deeper, evaluative peer-review roles for AI.

B. Contributions and Organization of the Paper

This work aims at contributing with the discussions concerning the use of AI during the peer review of scientific manuscripts. It attempts to reach this objective by:

- Presenting a structured framework to guide the use of AI in supporting peer review, clarifying distinct roles such as pre-review assistance and review enhancement.
- Assessing the technical capabilities and limitations of current AI tools in performing key review tasks.
- Identifying ethical and practical challenges associated with integrating AI into the peer review process.
- Providing illustrative use cases and offer guidelines for the responsible and effective use of AI by reviewers.

The remainder of the paper is organized as follows. Section II discusses the main advantages of using AI in manuscript review. Section III presents the key limitations and risks associated with AI-based evaluation. Section IV analyzes the distinction between novelty detection and genuine scientific contribution. Section V proposes a structured framework for AI-guided manuscript pre-evaluation, including expected inputs and outputs. Section VI introduces a prompt template designed to assist AI tools such as ChatGPT in evaluating manuscripts. Finally, Section VII offers concluding remarks and outlines directions for future research.

II. Advantages of AI-Based Review

This section discusses the primary advantages of integrating AI into the manuscript review process.

A. Speed and Scalability

One of the most immediate advantages of integrating AI into the manuscript review process lies in its potential to vastly increase throughput while significantly reducing latency. AI systems, particularly those based on natural language processing and document parsing, can analyze the structure, content, and metadata of a manuscript within seconds to minutes. This contrasts sharply with traditional peer review, where each review can take several hours of expert time, often spread over days or weeks due to competing obligations and communication delays.

In high-volume journals and large conference proceedings, where thousands of manuscripts are submitted annually, editorial teams face growing logistical pressures to assign reviewers, track deadlines, and ensure timely decision-making. AI tools can support this by automating the initial triage stage, quickly flagging submissions that are incomplete, improperly formatted, or clearly out of scope. This allows human reviewers and editors to focus on more substantive and complex cases, improving overall workflow efficiency [2].

Moreover, AI systems offer scalability that is virtually unbounded by human labor constraints. Once trained and validated, an AI model can operate continuously without fatigue, making it especially valuable for pre-evaluating large submission batches or managing rolling review schedules. This is particularly relevant for multidisciplinary mega-journals and open-access platforms that operate under continuous publication models.

Importantly, AI can also facilitate real-time analytics and feedback. For example, editorial dashboards can be augmented with AI-generated assessments of manuscript readability, citation integrity, or methodological soundness, enabling editors to make better-informed desk decisions. These capabilities help reduce backlog, accelerate the review cycle, and improve authors' experience by shortening time-to-decision metrics.

B. Consistency

A well-recognized limitation of traditional peer review is the variability in human assessments. Reviewer evaluations are influenced by numerous factors, including fatigue, time constraints, domain familiarity, personal biases, and even unconscious heuristics. These influences can lead to inconsistencies in how different reviewers assess the same manuscript, particularly when evaluating elements such as structure, formatting, referencing, and writing clarity [4]. Such variability is problematic not only for authors, who may receive conflicting feedback, but also for editors who must reconcile divergent opinions to make publication decisions.

Artificial Intelligence offers a pathway to greater uniformity in evaluation. By executing predefined procedures and scoring rules consistently across submissions, AI tools can ensure that each manuscript is subjected to the same initial scrutiny, irrespective of when or by whom the manuscript is processed. For example, tools such as *statcheck* [12] can automatically verify statistical reporting against common standards, while citation analysis engines can detect missing, outdated, or suspicious references with high reliability.

Furthermore, AI can enforce journal-specific guidelines at scale. Many editorial systems now integrate AI modules that check adherence to formatting rules, reference styles, section completeness, or ethical disclosures (e.g., funding and conflict of interest statements). These evaluations are not only repeatable but also free from the influence of cognitive load or environmental distractions that may compromise human consistency.

It is important to note, however, that consistency in AI is a function of the design of its rule set or training corpus. If the underlying data or algorithms encode bias, the uniformity of evaluation may amplify rather than mitigate disparities. Nonetheless, when used as a supplement to human review rather than a replacement, AI consistency enhances fairness and procedural transparency in routine editorial checks [2].

C. Technical Anomaly Detection

One of the most practical contributions of AI in the manuscript review process lies in its ability to detect technical anomalies at scale and with a level of consistency difficult to achieve with human reviewers alone. Advanced natural language processing (NLP), machine learning, and rule-based systems are increasingly capable of identifying issues such as plagiarism, improper citation practices, statistical inconsistencies, and even image manipulation, all of which are critical markers of scientific rigor and integrity.

Plagiarism detection is one of the most mature and widely adopted AI applications in publishing. Tools such as *Turnitin* and *iThenticate* [13], [14] leverage vast indexed corpora and machine learning algorithms to identify textual overlap, paraphrasing without citation, and duplication across submissions. These systems provide similarity reports with granular insights, allowing editors to distinguish between acceptable reuse (e.g., methods sections) and unethical copying.

In parallel, AI tools such as *Statcheck* [12] and the already-mentioned GRIM test are designed to flag statistical reporting inconsistencies. These systems automatically extract statistical test results from manuscripts and verify whether the reported p-values, means, and standard deviations are mathematically coherent and internally consistent. Empirical studies have shown that such tools can detect errors in up to 50% of published papers in psychology, many of which go unnoticed during manual review [5].

Another area of AI application is image forensics. Deep learning models trained on large datasets of scientific images can detect signs of duplication, manipulation, or inappropriate enhancement in figures, including western blots and microscopy images. Journals like *EMBO Reports* and *Molecular and Cellular Biology* have already implemented automated image screening in their editorial pipelines.

Citation integrity is another domain where AI excels. Algorithms can identify reference loops, suspicious citation clusters (e.g., citation cartels), and citations to retracted articles. These checks are especially valuable in detecting questionable scholarly practices and ensuring that manuscripts build upon sound and verified literature.

Although these tools are not infallible and sometimes produce false positives, their inclusion in editorial workflows enhances the precision and thoroughness of the review process. Importantly, they can serve as a first line of defense, flagging issues for further human investigation rather than making autonomous decisions, thus complementing rather than replacing the role of expert human reviewers.

D. Reviewer Assistance

While full automation of manuscript evaluation remains ethically and epistemologically problematic, the use of AI as an assistive tool for human reviewers presents a compelling case. AI technologies can streamline the review process by performing time-consuming but cognitively routine tasks, allowing reviewers to focus on the nuanced assessment of scientific validity, novelty, and impact.

One key area where AI proves helpful is in automatic summarization. Large language models (LLMs) and transformer-based architectures can generate concise summaries of manuscript content, identifying the main objectives, methods, and findings [15]–[17]. These summaries can be used as part of editorial dashboards or provided to human reviewers as a rapid orientation tool, especially for interdisciplinary submissions where the reviewer may not be familiar with all technical domains.

In addition to summarization, AI systems can highlight sections needing further scrutiny. For example, natural language processing tools can flag ambiguous phrasing, insufficient methodological detail, or unsupported conclusions. Grammar correction models can identify syntactic and semantic issues, while logic-checking tools can assist in detecting inconsistencies across sections of a manuscript (e.g., conclusions that do not align with stated claims or results).

Another crucial application lies in reviewer recommendation. AI systems trained on citation networks, topic modeling, and publication metadata can suggest reviewers based on their prior publications, citation overlap, or co-authorship networks. Tools such as Elsevier's *Reviewer Finder*, Springer's *Reviewer Recommender*, and Clarivate's *Publons* have already implemented reviewer matching algorithms based on machine learning [18]–[20]. This feature is especially valuable for early-career editors or when handling niche or interdisciplinary submissions for which expert reviewers are not immediately obvious.

Artificial intelligence can also support structured feedback generation. By aligning with journal-specific checklists or ethical guidelines, AI models can prompt reviewers to consider specific dimensions of quality such as reproducibility, data availability, and ethical approval declarations. This helps standardize review coverage and reduce the risk of important omissions.

These capabilities, when integrated into editorial systems, reduce the cognitive and administrative load on reviewers and editors alike, improving turnaround times without compromising review depth. However, the quality of AI-generated assistance depends on the training corpus and design of the underlying algorithms. Therefore, human oversight remains essential to validate, contextualize, and interpret AI suggestions appropriately.

E. Reduction of Reviewer Fatigue

Reviewer fatigue is a well-documented and growing concern within the scholarly publishing ecosystem. As submission volumes continue to rise across virtually all scientific fields, the availability of qualified reviewers has not kept pace, leading to increasing demands on a relatively static reviewer pool. Surveys and empirical studies have documented signs of reviewer overload, including delayed responses, superficial evaluations, and elevated rejection rates due to lack of detailed feedback.

Support systems based on AI can play a significant role in alleviating this burden by automating the more repetitive, time-consuming, and cognitively routine tasks in the review process. These include technical screening (e.g., checking for missing sections, mismatched citations, incorrect formatting), language and grammar correction, plagiarism detection, and basic statistical validation. By delegating these tasks to AI, reviewers are spared the mental load associated with non-substantive checks and can focus their cognitive effort on evaluating the originality, coherence, and scientific contribution of the manuscript.

This division of labor, where AI performs routine validation and human reviewers focus on interpretative judgment, embodies a form of hybrid intelligence. In such frameworks, the AI acts as a cognitive amplifier rather than a substitute, helping to increase reviewer throughput while preserving depth and attention to scientific nuance. Importantly, this approach also enhances reviewer engagement, as it aligns the review task with more meaningful and intellectually stimulating components of manuscript evaluation. It is stressed that such hybrid intelligence is the central claim of this work.

Several publishing platforms are already experimenting with integrated AI-assisted workflows that allow reviewers to receive preprocessed reports summarizing manuscript structure, ethical compliance flags, and reference quality assessments. For example, *Springer Nature* has piloted systems that provide automated initial checks on manuscripts prior to peer review, reporting improvements in overall review turnaround time and reviewer satisfaction [21].

Ultimately, mitigating reviewer fatigue through AI assistance has systemic benefits. It can help reduce delays in the peer review cycle, enhance the consistency of feedback, and prevent burnout in expert reviewers, many of whom serve voluntarily. As journals increasingly depend on this limited resource, sustainable reviewer engagement becomes essential to maintain the integrity of the academic publication process.

III. Disadvantages of AI-Based Review

This section discusses the primary disadvantages of integrating AI into manuscript review.

A. Inability to Assess Novelty and Context

A core limitation of current AI systems in the manuscript review process is their inability to evaluate the scientific novelty and contextual significance of research contributions, two fundamental dimensions that determine the merit of scholarly work. While AI models excel at pattern recognition, language modeling, and summarization based on vast training corpora, they inherently rely on what has already been published to infer patterns and assess content. Consequently, their outputs are bounded by the statistical regularities and dominant paradigms within existing literature.

This constraint implies that AI systems may fail to appreciate truly innovative, unconventional, or interdisciplinary research, especially when such work challenges prevailing norms or introduces unfamiliar methodologies. For example, a manuscript proposing a novel theoretical model or a breakthrough empirical finding may lack citation density or structural similarity with known works, features often used by AI systems for validation. As a result, the AI model may erroneously evaluate the paper as irrelevant, incomplete, or low quality.

Moreover, AI systems typically lack situated understanding, meaning they do not possess contextual awareness of how a specific piece of research fits within broader scientific debates, historical developments, or unresolved controversies. This is particularly problematic in emerging fields, where the literature base is sparse, or in socially sensitive domains, where ethical, cultural, and epistemological subtleties are central to evaluating contribution and impact.

Attempts to bridge this gap by fine-tuning models on specific domain corpora or introducing context-aware embeddings remain at early stages and still fall short of replicating the type of judgment exercised by expert reviewers who draw on tacit knowledge, disciplinary discourse, and years of experience. Furthermore, such efforts may even increase the risk of confirmation bias, as the model becomes increasingly tuned to prevailing conventions rather than disruptive insights.

In light of these constraints, relying on AI for triage or basic consistency checks may be justified, but using it as the primary or sole arbiter of scientific merit raises serious concerns. Human reviewers remain indispensable for detecting conceptual leaps, recognizing subtle theoretical advances, and appreciating the broader implications of a manuscript's claims. Therefore, any hybrid review model must clearly delimit the evaluative boundaries of AI systems and embed them within a human-in-the-loop framework.

B. Bias Reinforcement

A critical concern in the integration of AI into the manuscript review process is the risk of bias perpetuation. AI systems, particularly those based on machine learning and large language models, are inherently dependent on the datasets used for their training. When these datasets are drawn from existing scientific literature, they tend to reflect the structural inequalities and epistemic biases already present in scholarly publishing. Consequently, AI systems may inadvertently reinforce dominant paradigms, mainstream topics, and privileged institutional voices, undermining efforts to promote the epistemic pluralism that is welcome in science.

For example, several studies have shown that citation-based metrics, which are frequently used in AI-powered reviewer recommendation and manuscript ranking systems, disproportionately favor publications from English-speaking countries, high-impact journals, and elite academic institutions. When these metrics are embedded in AI decision-making processes, such as relevance scoring or quality prediction, they risk marginalizing novel or underrepresented contributions, particularly from early-career scholars, or interdisciplinary authors whose work does not yet align with established citation patterns.

Topic modeling and similarity scoring can also exacerbate thematic conservatism. AI tools designed to flag 'out-of-scope' or 'low-relevance' submissions may implicitly penalize research that explores emerging topics, challenges prevailing theories, or adopts unconventional methodologies. This algorithmic conservatism can have chilling effects on creativity and innovation, pushing authors to conform to mainstream scientific discourse in order to meet AI-driven screening thresholds.

Moreover, while some have raised concerns about bias and underrepresentation in automated systems, AI-based manuscript evaluation tools are not inherently (and should not be) responsible for correcting disparities in publication patterns across gender, geography, or institutional prestige. Their primary function is to assist with objective, content and merit-based analysis, such as structural consistency, clarity, or technical soundness, rather than adjudicate socio-political equity. Nonetheless, transparency in how such systems are trained and how decisions are generated remains important to ensure accountability and to avoid unintended distortions in the review process [7].

To mitigate the above-mentioned risks, the deployment of AI in manuscript evaluation must be accompanied by bias auditing protocols, transparency standards, and human-in-the-loop governance structures that prioritize fairness and accountability. Incorporating diverse training datasets, adopting interpretable models, and ensuring editorial oversight are essential safeguards to prevent AI from amplifying systemic inequities in science.

C. Lack of Ethical Judgment

A fundamental limitation of current AI systems in manuscript evaluation lies in their inability to exercise moral reasoning or make context-sensitive ethical judgments. While AI can detect certain forms of misconduct, such as plagiarism or image duplication, through pattern recognition and rule-based checks, it cannot evaluate deeper dimensions of ethical integrity, such as responsible conduct of research, social harm potential, conflicts of interest, or the broader societal implications of scientific claims.

Scientific peer review is not merely a technical filtering process; it is also an epistemic and ethical practice. Human reviewers are expected to weigh questions of fairness, transparency, and accountability. For instance, determining whether a study involving human participants has obtained valid ethical clearance, or whether it risks exacerbating social inequities (e.g., through biased datasets or harmful inferences), depends on cultural context, disciplinary norms, and value-based reasoning, dimensions that current AI systems are not equipped to process. These forms of interpretive judgment require human expertise, as they extend beyond the computational capabilities of existing AI models.

Moreover, ethical evaluations often require balancing competing interests. A study may be methodologically sound yet raise ethical concerns due to its potential applications (e.g., surveillance technologies, dual-use research). These value-laden trade-offs cannot be resolved by AI systems trained to optimize predictive accuracy. Instead, they require dialogue, dissent, and reflexivity, hallmarks of human ethical deliberation that resist algorithmic formalization.

The inability of AI to reason about intentionality, motivation, or long-term societal impact becomes particularly problematic when reviewing research in areas such as biomedicine, climate science, social policy, or even artificial intelligence itself. In such fields, ethical considerations are inseparable from scientific merit. For example, a paper proposing a novel genetic modification technique may be technically valid, but still ethically contentious. Human reviewers, informed by codes of conduct like the Committee on Publication Ethics (COPE), the World Medical Association (WMA), and the Belmont Report [22]–[24] are needed to contextualize such work within responsible research and innovation frameworks.

Efforts to embed ethical reasoning in AI, such as integrating rules from bioethics or philosophy into decision trees, remain limited, domain-specific, and brittle. Attempts to train models on annotated ethical judgments also face significant challenges due to the lack of consensus on ethical norms and the difficulty of operationalizing complex ethical questions.

Given these constraints, AI should not be entrusted with independent ethical adjudication in peer review. Instead, AI-generated outputs must be explicitly flagged as ethically incomplete, with the understanding that final evaluations must rest with human reviewers trained in the ethical dimensions of scholarly publishing.

D. Risk of Dehumanization

Integrating AI into the peer review process also introduces concerns about the potential dehumanization of scholarly evaluation. While AI can enhance efficiency, its excessive dependence may compromise the human elements essential to academic discourse, such as empathy, ethical judgment, and nuanced understanding.

Research indicates that interactions with AI systems can lead to perceptions of dehumanization. A study by Dang and Liu [25] categorizes AI-induced dehumanization into three levels: human-AI interaction, intrapersonal dynamics, and interpersonal relationships. They highlight that users may experience diminished self-regard and social relationships when engaging extensively with AI systems.

In the context of peer review, the absence of human judgment in AI assessments can result in the neglect of ethical considerations and the broader societal implications of research. Doskaliuk et al. [26] emphasize that while AI tools can assist in preliminary evaluations, they lack the capacity to fully comprehend complex scientific content and ethical nuances, necessitating human oversight to maintain the integrity of the review process. Moreover, the anthropomorphism of AI systems can lead users to attribute human-like qualities to machines, potentially eroding the authenticity of human interactions.

Realizing the benefits of AI-assisted editorial processes requires careful and deliberate system design. Poorly calibrated AI pipelines risk either inundating editors with false positives or, conversely, allowing unsuitable submissions to bypass scrutiny. Therefore, scalability must be balanced with accuracy, transparency, and well-defined thresholds to ensure that AI serves as a reliable triage tool rather than a source of additional editorial burden. Furthermore, editorial judgment is nuanced. A manuscript may exhibit several weaknesses, yet an experienced editor might choose to advance it through the review process due to a particularly compelling or innovative contribution. Conversely, a single critical flaw may warrant rejection, even if the remainder of the manuscript is of high quality. In contrast, AI systems tend to evaluate all criteria uniformly, potentially failing to detect issues that seasoned editors would recognize as decisive.

To mitigate these risks, it is crucial to maintain a balanced integration of AI in peer review, ensuring that human reviewers remain central to the evaluation process. This approach preserves the ethical standards and humanistic values fundamental to scholarly communication.

E. *Susceptibility to Gaming*

The incorporation of AI into the manuscript evaluation pipeline introduces new opportunities for manipulation. As AI systems become more predictable in their scoring patterns or linguistic preferences, authors may learn how to craft submissions that align with these preferences without genuinely improving the quality of their research.

This phenomenon, often referred to as *gaming the system*, has precedent in academic publishing. Similar effects have been observed in journal ranking and citation metrics, where researchers have adjusted behavior to satisfy algorithmic evaluation tools [27]. In the context of AI peer review, this might include artificially inflated clarity, the use of keywords known to trigger positive assessments, or even the inclusion of self-citations to favorably influence AI-based citation analysis.

Unlike human reviewers, AI systems generally lack the contextual awareness to detect when surface-level compliance masks deeper methodological or conceptual flaws. As a result, manuscripts that are formally well-structured but scientifically weak may be rated more favorably by automated evaluators than by critical human experts. This risk is particularly pronounced if AI systems are deployed without sufficient oversight or transparency.

To minimize susceptibility to gaming, peer review frameworks must treat AI assessments as assistive rather than definitive. Furthermore, transparency in evaluation criteria and regular auditing of AI outputs are essential to prevent exploitation of algorithmic blind spots.

F. *Lack of Transparency and Accountability*

The integration of AI into the peer review process also introduces concerns regarding transparency and accountability. AI systems, particularly those utilizing complex machine learning algorithms, often operate as ’black boxes’, making it challenging to understand the rationale behind their decisions. This opacity can undermine trust in the peer review process and complicate the identification and correction of errors. In human review, accountability is typically linked to named or anonymous reviewers and editors. In AI-assisted review, concerns arise because automated systems may make impactful decisions without transparent reasoning or clear assignment of responsibility.

Cheong [28] emphasizes that the lack of transparency in AI decision-making processes poses ethical and legal challenges, especially when these systems impact individual and societal well-being. Because many AI systems operate as black boxes, their outputs may be difficult to interpret or justify. This opacity can result in decisions that are not only unexplainable but may also reflect hidden biases embedded in the training data. Such limitations raise legitimate concerns about transparency and accountability, particularly in contexts like scholarly publishing, where evaluative decisions must be both fair and defensible.

In the context of academic peer review, the American Society for Microbiology (ASM) highlights that current AI tools do not meet the transparency and accountability standards required for scholarly publishing. The ASM notes that while AI can assist in the review process, it cannot replace the nuanced judgment of human reviewers, and its use must be carefully managed to maintain the integrity of the review process [29].

Hence, it is crucial to implement AI systems that are explainable and to establish clear guidelines for their use in peer review. This includes ensuring that AI tools are used to support, rather than replace, human judgment and that their decision-making processes are transparent and accountable.

G. *Confidentiality and Legal Concerns*

Peer review is a confidential process, and uploading manuscripts to publicly accessible AI tools may violate the trust inherent in this system and breach confidentiality agreements. Such use is only permissible when the journal has an explicit policy allowing authors to consent to AI-assisted peer review, and when authors have actively opted in to this arrangement [30]. In the absence of such provisions, reviewers are expected to assess manuscripts using their own expertise, without reliance on AI tools. The undisclosed use of AI may compromise transparency and constitute a violation of academic integrity norms.

From a legal standpoint, inputting another person's unpublished manuscript into an AI system may infringe upon copyright laws and breach data protection regulations. During the peer review process, journals or publishers often hold the rights to the manuscript, and unauthorized distribution of this content may constitute copyright infringement. Moreover, many AI platforms store and process user inputs, which may contravene data privacy laws, such as the General Data Protection Regulation (GDPR) [31], particularly when sensitive or identifiable information is involved. Such actions may also breach the terms of service of either the AI provider or the publisher, exposing reviewers and their affiliated institutions to legal and reputational risks.

IV. Novelty Detection vs Scientific Contribution

While AI systems can detect textual and semantic novelty by comparing manuscript content with previously published literature, they are fundamentally limited in assessing genuine scientific contribution.

Modern AI systems, especially those based on large language models and retrieval-augmented generation (RAG), are proficient in identifying novelty in the semantic and lexical sense [32]. That is, AI can: i) compare claims made in the manuscript against vast corpora of existing literature; ii) detect linguistic and conceptual deviations from known patterns, formulas, algorithms, or conclusions; iii) flag sections that appear to introduce methods, results, or hypotheses not found in previously indexed or trained data.

This enables AI to estimate relative novelty, i.e., whether a claim appears new with respect to documented knowledge. However, and very relevant, this novelty is syntactic and relational, not epistemic. The AI determines what is unfamiliar, not necessarily what is important. In other words, such assessments are based on surface features and past training data. They indicate what is unfamiliar, not necessarily what is valuable.

A. Limits in Assessing Scientific Contribution

The evaluation of scientific contribution extends well beyond the detection of surface-level novelty. While AI systems, especially large language models, are increasingly capable of identifying unusual phrasing, new terminology, or deviations from dominant research patterns, they fall short of grasping the deeper intellectual and epistemological value that characterizes genuine scientific progress. Assessing contribution involves a multidimensional judgment that encompasses both the internal rigor of the work and its broader impact within and beyond the research field.

One critical dimension is the theoretical significance of the research. A contribution may be considered meaningful if it resolves a longstanding problem, proposes a synthesis across previously disconnected theoretical frameworks, or introduces a novel conceptual model that has the potential to reorganize the landscape of inquiry. Such achievements often depend on insight into the historical evolution of ideas, the recognition of subtle gaps in understanding, and an appreciation for the intellectual boldness required to question entrenched assumptions, all of which require deep disciplinary familiarity and interpretative capacity.

Another essential consideration is the practical utility of the findings. Contributions that lead to more efficient algorithms, more accurate measurement techniques, or actionable insights for real-world systems, be they technological, social, economic, or environmental, often indicate scientific maturity and relevance. Yet practical utility alone is insufficient; transformative contributions typically also possess explanatory power and raise new, generative questions for future inquiry.

Conceptual innovation is equally vital. Some manuscripts may appear modest in scope but introduce a redefinition of key constructs or a reinterpretation of established phenomena. These shifts, while subtle, can be epistemologically profound and lead to cumulative change. Conversely, work that appears lexically novel may simply rearrange well-known components without advancing understanding, a distinction that current AI systems, focused primarily on textual surface patterns and citation statistics, are not yet equipped to discern reliably.

A final but fundamental aspect is the generalizability and potential long-term influence of the work. Contributions that offer robust findings across datasets, settings, or methodologies tend to shape research agendas and influence scholarly discourse. Recognizing such potential requires not only technical literacy but also predictive judgment, an ability to anticipate how a piece of research might catalyze subsequent developments. This is inherently forward-looking and context-sensitive, two dimensions where human reviewers, informed by experience and disciplinary dialogue, retain an advantage relative to AI systems.

Hence, evaluating scientific contribution requires the integration of theoretical depth, methodological soundness, conceptual novelty, practical impact, and visionary judgment. These are not simply computational challenges; they are epistemic tasks that demand philosophical insight, familiarity with the evolving norms of scientific discourse, and an ability to situate a work within a broader intellectual trajectory.

Because AI systems are trained exclusively on past data and lack intentionality, historical awareness, and value-based reasoning, their capacity to assess scientific contribution remains fundamentally limited. At best, AI can serve as a filter or aid in preliminary screening. The authoritative judgment on what truly advances science must remain the responsibility of expert human reviewers.

B. Illustrative Contrast and Implications for Peer Review

The distinction between novelty and true scientific contribution becomes particularly evident when we consider practical examples. Suppose a manuscript introduces a new quantum error-correction scheme that, on the surface, shares substantial linguistic similarity with existing schemes. Despite this superficial overlap, it proposes a fundamentally new encoding logic that represents a genuine paradigm shift. An AI system, heavily reliant on pattern matching and lexical similarity, may undervalue the submission, categorizing it as lacking in novelty. Conversely, a manuscript that rearranges well-known components using distinct phrasing or unconventional sectioning might be flagged as highly novel by the same AI model, despite offering little in terms of substantive scientific advancement. The section Susceptibility to Gaming is closely related to this example.

Such cases underscore the necessity of reserving interpretive and epistemic authority for human experts. While AI can provide valuable assistance, such as identifying overlapping text, detecting formulaic patterns, or flagging anomalies, its evaluations must be understood as preliminary and subject to human review. AI systems are capable of detecting textual novelty, highlighting deviations from established patterns, and suggesting areas for closer scrutiny. However, they are not equipped to determine whether a perceived novelty is scientifically meaningful, theoretically significant, or contextually transformative.

This reinforces the view that AI should function as a support tool in peer review rather than a substitute for expert judgment. Human reviewers bring disciplinary insight, historical context, and philosophical grounding to the task of manuscript evaluation. They can assess whether a contribution resolves an open problem, shifts a theoretical framework, or offers long-term impact across domains, dimensions that remain beyond the reach of current AI capabilities.

Table I synthesizes a comparison of AI capabilities for manuscript review.

TABLE I
COMPARISON OF AI CAPABILITIES IN MANUSCRIPT REVIEW

Aspect	AI capability	Limitation
Detecting textual novelty	Strong	Cannot distinguish deep vs. superficial novelty
Assessing scientific value	Weak	Cannot contextualize impact or generalizability
Supporting peer review	Useful pre-filter	Must defer final evaluation to human reviewers

V. AI-Guided Manuscript Pre-Evaluation: Input Topics and Expected Outputs

To support a fair, context-aware, and efficient pre-evaluation process, an AI-assisted manuscript review system must operate on a well-defined set of structured inputs. These include the manuscript file itself, along with metadata and contextual information that enable a nuanced interpretation of content and alignment with journal or conference standards. It is important to stress that the role of the AI is not to replace expert reviewers, but to provide a preliminary assessment that streamlines the review workflow and enhances consistency across submissions.

Among the essential contextual inputs, the system should receive the manuscript's title and abstract to identify the central topic and assign it a preliminary thematic classification. Author information, although excluded in blinded reviews, may be used for identifying citation self-references or conflicts of interest when available. The declared target journal or conference scope enables the system to assess venue relevance, while the declared manuscript type, such as research article, short communication, or case study, guides format and expectation alignment. Furthermore, access to the venue's reviewer guidelines or evaluation criteria allows the AI system to align its assessment with specific editorial priorities such as novelty, technical soundness, clarity, or reproducibility.

Once the manuscript and context are processed, the AI system should analyze the full content across several key dimensions. First, structural and formal aspects must be verified, including the presence and formatting of title pages, abstracts, keywords, sections, and references. The AI should also evaluate language clarity, grammatical correctness, text coherence, redundancy, and adherence to expected length constraints.

Second, the system should assess novelty by identifying explicitly claimed contributions and comparing semantic similarity with the existing literature to estimate originality. Although AI cannot make epistemological judgments, it can flag passages where novelty is claimed and estimate lexical uniqueness relative to published corpora.

Third, technical soundness must be evaluated through the verification of logical argumentation, statistical validity, mathematical derivations, model assumptions, and any code or algorithmic artifacts that are included.

Fourth, the AI should assess experimental design and results presentation, verifying whether datasets are well-described, whether evaluation metrics and baselines are appropriate, whether results are statistically robust, and whether stated hypotheses align with conclusions. Reproducibility is a major concern, and the AI must be able to check for the presence of open data, code availability, and replicability of methodology.

Fifth, the system should evaluate the related work section by checking for completeness, identifying missing foundational or contemporary references, and detecting irregular citation behaviors such as excessive self-citation or irrelevant references. Sixth, ethical compliance must also be reviewed. This includes verifying the presence of ethics approval statements for human or animal data, availability of data and code, and scanning for signs of manipulated images or fabricated content.

Based on this comprehensive analysis, the AI system should generate outputs that assist the human reviewer. These include a concise manuscript summary in abstract style, accompanied by bullet-point highlights of the claimed contributions.

It should also produce a preliminary strengths-and-weaknesses summary, referencing specific sections or figures as evidence. Additional outputs should include an assessment of whether the manuscript fits the journal's scope, a novelty estimate based on semantic overlap with prior literature, and alerts regarding technical inconsistencies or missing assumptions.

A reproducibility checklist should summarize the status of data/code availability and methodological transparency. Ethical and compliance flags may be raised when critical declarations are missing or when statistical patterns appear anomalous. In areas of ambiguity or limited interpretability,

Prompt 1 for manuscript review with ChatGPT

Please, do a rigorous review of the attached manuscript. Prepare a review report containing a recommendation regarding the suitability of the manuscript for publication, as well as recommendations to the authors if clear opportunities for improvements are identified. Your output must follow these guidelines:

1) Concise summary and main contributions: Write a 1-paragraph summary of the manuscript, followed by a bullet list of the main contributions.
2) Relevance: Evaluate the alignment of the manuscript with the target journal scope: [insert journal/conference scope here].
3) Novelty assessment based on claimed contributions: Identify the main claims of novelty. If possible, estimate their originality based on similarity to existing literature.
4) Technical soundness and logic consistency: Highlight potential issues in logic, theoretical derivations, experimental design, or statistical reasoning.
5) Clarity and structure: Comment on the organization, clarity of language, and compliance with scientific writing conventions.
6) Reproducibility: Check whether code, datasets, or sufficient methodological detail is provided to enable replication.
7) Ethical concerns and research integrity: Flag any ethical or integrity-related concerns (e.g., missing declarations, image manipulation, or unverifiable claims).
8) Coverage and correctness of references: Assess whether the related work section is comprehensive and whether important references are missing or misused.
9) Questions to guide human reviewers: Propose specific questions or points that a human reviewer should pay special attention to during their evaluation.
10) Optional (if available): Perform figure/table quality analysis and flag resolution, labeling, or consistency issues.
11) Optional (if available): Evaluate for potential demographic or methodological bias (if human/social data is involved).

the system should generate reviewer-specific questions to guide deeper human analysis.

Finally, optional advanced analyses may include citation network profiling to detect citation loops or insular author clusters, image and table quality checks to evaluate resolution and labeling standards, and bias detection models in papers using human or social datasets. These outputs, while not definitive, can highlight areas for deeper expert scrutiny and improve the consistency and fairness of the peer review process.

VI. Templates for AI-Assisted Pre-Evaluation and Author Testimony

This section introduces two structured prompt templates designed for use with AI tools such as ChatGPT to support the pre-evaluation of scientific manuscripts. These templates aim to elicit feedback aligned with conventional peer review criteria and can serve as a preparatory step before review report submission. The section also includes the author's testimony, offering reflections on the application of these prompts and observations regarding recent trends in AI-influenced peer review reports.

A. Prompt Templates for ChatGPT Evaluation

Prompt 1 offers a detailed and methodically structured query intended for submission to ChatGPT alongside the manuscript file. This prompt is designed to initiate a comprehensive AI-supported pre-evaluation, aligned with key assessment dimensions discussed earlier in the manuscript. Users are advised to adapt point 2 by specifying the scope of the target journal or conference to ensure contextual relevance.

Prompt 2 presents an alternative review framework inspired by the guidelines used by the Journal of Communications and Information Systems (JCIS), under the *Sociedade Brasileira de Telecomunicações* (SBrT). These criteria are influenced by peer review standards from *Nature Communications* [9], emphasizing clarity, rigor, and ethical responsibility in manuscript evaluation. The prompt reflects a more question-driven approach, helping reviewers focus on substantive aspects of the manuscript review process.

B. My Testimony as an Author

As a researcher accustomed to submitting manuscripts to scientific journals and engaging with peer feedback, I have increasingly observed signs suggesting the possible use of AI not as a supportive tool for reviewers, but as a substitute for them. This shift appears to correlate with a rise in review reports that resemble extensive to-do lists, often misaligned with the core focus of the manuscript. Rather than offering targeted, constructive critique, some reviews extend well beyond the central contribution, presenting demands that are, at best, difficult to implement and, at worst, practically unfeasible.

These requests frequently aim to broaden the manuscript's scope far beyond its original objectives. If fully addressed, they would risk transforming the manuscript into a disjointed assemblage of loosely connected additions, more akin to a patchwork quilt than a coherent scholarly narrative. Such revisions compromise clarity and focus, and may ultimately cause the manuscript to exceed the journal's length restrictions. This trend raises concerns not only about the quality and fairness of the peer review process, but also about the consequences of strongly relying on AI-generated evaluations without sufficient human oversight.

On the authoring side of scientific research, I have integrated AI tools into my workflow as a valuable resource for manuscript pre-evaluation. Prior to submission, I routinely employ language models to obtain preliminary feedback on structure, clarity, methodological rigor, and compliance with editorial standards. Although the outputs occasionally exhibit some of the limitations discussed earlier, such as superficial novelty detection or mechanical suggestions, the insights they offer often help to streamline the manuscript, eliminate ambiguities, and ensure internal coherence. In particular, the use

Prompt 2 for manuscript review with ChatGPT

Please, do a rigorous review of the attached manuscript. Prepare a review report containing a recommendation regarding the suitability of the manuscript for publication, as well as recommendations to the authors if clear opportunities for improvements are identified. Your output must answer the following questions:

1) Major claims and significance: What are the major claims of the paper? and Will the paper be of interest to others in the field?
2) Novelty: Are the claims novel? If not, please identify the major papers that compromise novelty.
3) Convincing claims: Are the claims convincing? If not, what further evidence is needed?
4) Contextual discussion: Are the claims appropriately discussed in the context of the previous literature?
5) Methodological details: Evaluate whether the methodology is sufficiently detailed to allow reproduction of the experiments and results.
6) Suggestions for improvement: Suggest additional experiments or data that could strengthen the work.
7) Clarity and presentation: Comment on the clarity and accessibility of the manuscript, including the abstract, introduction, and conclusions.
8) Supplementary materials: Consider whether authors should provide supplementary methods or data, such as source code or detailed protocols.
9) Balanced presentation: Assess whether the authors have presented their claims without overselling and have treated previous literature fairly.
10) Potential for resubmission: If the manuscript is deemed unacceptable, consider whether the study is promising enough to warrant resubmission and specify what work is needed to make it acceptable.
11) Impact of further work: Evaluate how much further work would improve the contribution, how difficult this would be, and how long it might take.

of the prompt templates presented in this section has proven effective in simulating the types of questions and critiques that might emerge during formal peer review, allowing for proactive refinement of the manuscript before external evaluation

VII. Conclusions

This paper explored the feasibility, limitations, and implications of using AI tools in scientific peer review processes. The central claim is that AI tools can significantly augment or partially replace human peer review tasks, with potential gains in efficiency and consistency but also possible risks regarding epistemological integrity and ethical considerations.

The integration of AI into the manuscript review process holds potential for improving speed, consistency, and efficiency. However, full automation of scientific peer review is not advisable at this stage due to AI's current limitations in handling conceptual, ethical, and community-based aspects of research evaluation.

The most promising approach appears to be a hybrid model, where AI acts as a pre-screening and assistance tool, while final judgments and nuanced critiques remain under the domain of human experts. This preserves the strengths of both systems and maintains the integrity of the scholarly process.

References

[1] S. P. J. M. Horbach and W. Halffman, "Journal peer review and editorial evaluation: Cautious innovator or sleepy giant?" *Minerva*, vol. 58, no. 2, pp. 139–161, 2020.

[2] D. Heaven, "AI peer reviewers unleashed to ease publishing grind," *Nature*, vol. 563, no. 7733, pp. 609–610, 2018.

[3] F. Squazzoni, "Editorial note: We need to recognise that peer review is central to the 'social contract' of academic citizenship," *Journal of Artificial Societies and Social Simulation*, vol. 28, no. 1, p. 6, 2025.

[4] C. J. Lee, C. R. Sugimoto, G. Zhang, and B. Cronin, "Bias in peer review," *Journal of the American Society for Information Science and Technology*, vol. 64, no. 1, pp. 2–17, 2013.

[5] N. J. L. Brown and J. A. J. Heathers, "The GRIM test: A simple technique detects numerous anomalies in the reporting of results in psychology," *Social Psychological and Personality Science*, vol. 8, no. 4, pp. 363–369, 2017.

[6] M. J. Mrowinski, P. Fronczak, A. Fronczak, M. Ausloos, and O. Nedic, "Artificial intelligence in peer review: How can evolutionary computation support journal editors?" *PLOS ONE*, vol. 12, no. 9, p. e0184711, 2017.

[7] M. Mitchell, S. Wu, A. Zaldivar, P. Barnes, L. Vasserman, B. Hutchinson, E. Spitzer, I. Raji, and T. Gebru, "Model cards for model reporting," in *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19)*, 2019, pp. 220–229.

[8] B. Doskaliuk, O. Zimba, M. Yessirkepov, I. Klishch, and R. Yatsyshyn, "Artificial intelligence in peer review: Enhancing efficiency while preserving integrity," *Journal of Korean Medical Science*, vol. 40, no. e92, 2025.

[9] Nature Communications, "For reviewers," 2025. [Online]. Available: https://www.nature.com/ncomms/for-reviewers

[10] M. J. Mrowiński, P. Fronczak, A. Fronczak, M. Ausloos, and O. Nedić, "Artificial intelligence in peer review: How can evolutionary computation support journal editors?" *PLOS ONE*, vol. 12, no. 9, p. e0184711, 2017. [Online]. Available: https://doi.org/10.1371/journal.pone.0184711

[11] A. Checco, L. Bracciale, M. Loreti, M. Petrocchi, and A. Spognardi, "AI-enabled manuscript screening: Risks, benefits, and limitations," *Learned Publishing*, vol. 34, no. 1, pp. 14–23, 2021. [Online]. Available: https://doi.org/10.1002/leap.1340

[12] M. Nuijten and J. Polanin, "Statcheck: Automatically detect statistical reporting inconsistencies to increase reproducibility of meta-analyses," *Research Synthesis Methods*, vol. 11, no. 5, pp. 574–579, 2020.

[13] Turnitin, LLC, "Turnitin similarity," 2025. [Online]. Available: https://www.turnitin.com/products/similarity/

[14] ——, "iThenticate: Plagiarism detection software," 2025. [Online]. Available: https://www.ithenticate.com/

[15] S. Minaee, T. Mikolov, N. Nikzad, M. Chenaghlu, R. Socher, X. Amatriain, and J. Gao, "Large language models: A survey," *arXiv preprint arXiv:2402.06196*, 2024. [Online]. Available: https://arxiv.org/abs/2402.06196

[16] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017. [Online]. Available: https://arxiv.org/abs/1706.03762

[17] H. Naveed, A. U. Khan, S. Qiu, M. Saqib, S. Anwar, M. Usman, N. Barnes, and A. Mian, "A comprehensive overview of large language models," *arXiv preprint arXiv:2307.06435*, 2023. [Online]. Available: https://arxiv.org/abs/2307.06435

[18] Elsevier, "Find reviewers tool," 2023. [Online]. Available: https://www.elsevier.com/connect/recap-and-recent-enhancements-to-find-reviewers

[19] Springer Nature, "Reviewer finder," 2023. [Online]. Available: https://www.springernature.com/gp/editors/resources-tools/reviewer-finder

[20] Clarivate, "Publons reviewer connect," 2018.

[21] Springer Nature, "Springer nature launches ai tool to help streamline integrity and ethics checks in publishing," 2024. [Online]. Available: https://www.springernature.com/gp/group/media/press-releases/ai-tool-to-help-streamline-integrity-and-ethics-checks/27730892

[22] Committee on Publication Ethics, "COPE code of conduct," 2008. [Online]. Available: https://publicationethics.org/files/2008%20Code%20of%20Conduct.pdf

[23] World Medical Association, "WMA international code of medical ethics," 2022. [Online]. Available: https://www.wma.net/policies-post/wma-international-code-of-medical-ethics

[24] National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, "The belmont report: Ethical principles and guidelines for the protection of human subjects of research," 1979. [Online]. Available: https://www.hhs.gov/ohrp/regulations-and-policy/belmont-report/index.html

[25] J. Dang and L. Liu, "Dehumanization risks associated with artificial intelligence use," *Journal of Personality and Social Psychology*, 2024.

[26] B. Doskaliuk, O. Zimba, M. Yessirkepov, I. Klishch, and R. Yatsyshyn, "Artificial intelligence in peer review: Enhancing efficiency while preserving integrity," *Journal of Korean Medical Science*, vol. 40, no. 7, p. e92, 2025.

[27] D. Moher, L. Bouter, S. Kleinert, P. Glasziou, M. H. Sham, V. Barbour, A. Cambon-Thomsen, A. M. Gülmezoglu, D. Hren, P. Glasziou *et al.*, "Assessing scientists for hiring, promotion, and tenure," *PLOS Biology*, vol. 16, no. 3, p. e2004089, 2018.

[28] B. C. Cheong, "Transparency and accountability in AI systems: safeguarding wellbeing in the age of algorithmic decision-making," *Frontiers in Human Dynamics*, vol. 6, p. 1421273, 2024.

[29] American Society for Microbiology, "AI in peer review: A recipe for disaster or success?" 2024, accessed: 2025-06-02. [Online]. Available: https://asm.org/articles/2024/november/ai-peer-review-recipe-disaster-success

[30] T. I. Leung, T. de Azevedo Cardoso, A. Mavragani, and G. Eysenbach, "Best practices for using ai tools as an author, peer reviewer, or editor," *Journal of Medical Internet Research*, vol. 25, p. e51584, Aug 2023.

[31] "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," https://eur-lex.europa.eu/eli/reg/2016/679/oj, Apr 2016, official Journal of the European Union, L 119, 4 May 2016, pp. 1–88.

[32] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, H. Küttler, M. Lewis, W.-t. Yih, T. Rocktäschel, S. Riedel, and D. Kiela, "Retrieval-augmented generation for knowledge-intensive nlp tasks," in *Advances in Neural Information Processing Systems*, vol. 33, 2020, pp. 9459–9474. [Online]. Available: https://arxiv.org/abs/2005.11401

Dayan Adionel Guimarães received his Master and Ph.D. degrees in Electrical Engineering from the State University of Campinas (Unicamp), Brazil, in 1998 and 2003, respectively. He is a Researcher and Senior Lecturer at the National Institute of Telecommunications (Inatel) in Brazil. His research is currently directed towards wireless communications in general, specifically radio signal propagation, digital transmission, spectrum sensing, dynamic spectrum access, and random signal processing.