Mukesh Kumar

4 September 2026

Synthetic Hermeneutics: A Protocol for Reading Literature in the Age of Generative AI

Abstract

Generative AI changes the objects that literary criticism is asked to interpret. A poem, novel, translation, or interactive narrative may now be shaped by prompts, model versions, training archives, interfaces, editorial selection, platform policies, and disclosures that reach readers at different moments. Existing approaches offer important resources for this situation, but no single approach reliably joins close reading to process evidence, archival uncertainty, material conditions, and institutional responsibility. This article proposes Synthetic Hermeneutics as a bounded protocol for that task. It defines a synthetic text as a text substantially generated, transformed, translated, structured, or personalized through machine-learning systems, and treats the published work as a synthetic textual event rather than as language detached from its conditions of production and reception. The method proceeds through five ordered readings: textual, production, archive, infrastructure, and responsibility. Each reading specifies its permissible evidence and its prohibited inferences. The protocol is illustrated through published cases including Ross Goodwin's *1 the Road*, K Allado-McDowell's *Pharmako-AI*, Jukka Aalho's *Aum Golly* projects, and Stephen Marche's *Death of an Author*. The article argues that literary criticism should neither infer hidden machine use from style nor abandon interpretation when records are incomplete. Instead, critics should state what is documented, what is attributed, what is inferred, and what remains unknown. This evidence-bound practice makes room for literary judgment while resisting claims that exceed the available record.

Keywords: generative AI; literary theory; hermeneutics; provenance; authorship; translation; infrastructure; disclosure

Introduction

Generative AI has made literary interpretation more difficult in a specific way. The difficulty is not simply that a model can produce prose. Literary works have always depended on distributed activity: writers, editors, translators, printers, publishers, readers, and technical systems all participate in the production and circulation of texts. What has changed is the density and instability of the mediations that may become relevant to the work. A language model can generate candidate passages; an editor can select, recombine, and rewrite them; a provider can alter a model version without altering the work's title; a platform can personalize the text presented to a reader; and a publisher can disclose machine involvement before, after, or never during the act of reading. These conditions do not replace literary form. They alter the evidentiary situation in which form is interpreted.

The familiar question, "Did AI write this?", is therefore inadequate. It compresses several distinct questions: What language or structure entered the published object? Which humans and systems participated in the workflow? What source materials and cultural archives made the system possible? Which interface, platform, labor, energy, and access conditions shaped the work's circulation? Who has the capacity to disclose, correct, compensate, or prevent harm? A single answer to authorship cannot settle all of these questions. Nor can a polished close reading settle them by itself.

This article proposes Synthetic Hermeneutics as an evidence-governed method for reading machine-learning-mediated literature. The proposal draws on philosophical hermeneutics, genetic criticism, book history, media archaeology, digital and computational

hermeneutics, and scholarship on AI's material and institutional conditions. It is not offered as a replacement for those traditions or as a claim that this framework is the first response to AI and interpretation. Recent work in computational hermeneutics, for example, treats generative AI as a cultural technology that should be evaluated through humanistic inquiry rather than accepted as a neutral instrument (Kommers et al.). The present proposal has a narrower task: to give literary critics a repeatable sequence for connecting a named textual object to a documented and partly undocumented field of mediation.

The method has two commitments. First, interpretation remains indispensable. A text may establish a voice, rhythm, genre relation, ethical address, or formal tension even when its production archive is incomplete. Second, textual appearance cannot prove a hidden workflow. Smooth prose, recurrent syntax, a detector score, or resemblance to a familiar author can support questions and comparisons; they do not independently identify a model, recover a prompt, establish training membership, or calculate a percentage of machine contribution. The critic must keep literary inference and causal assertion distinct.

The argument proceeds in four stages. It first defines the synthetic textual object and explains why a fixed percentage cannot decide the category. It then introduces the provenance field and the five readings that organize the protocol. The third section applies the method's claim discipline to style, translation, archives, material conditions, and disclosure. The final section shows why authorship, copyrightability, ownership, and responsibility must remain separate, especially when institutions rely on automated detection. Throughout, the purpose is not to make criticism procedural for its own sake. It is to make the grounds and limits of a critical claim visible.

From Synthetic Text to Synthetic Textual Event

The term *synthetic text* names a practical classification rather than an ontological verdict. A text qualifies when machine-learning systems substantially generate, transform, translate, structure, or personalize it. The criterion is dependency, not a percentage. A model may supply only a small number of words yet materially determine a work's plot constraint, translation route, recurrent structure, or personalized realization. Conversely, a large volume of generated material may become incidental if a later process removes every consequential trace from the published object. The critical question is whether the machine-learning-mediated operation remains constitutive of the work being interpreted.

This definition avoids two opposite mistakes. The first treats every software-assisted work as AI literature. Spellcheck, file storage, or ordinary word processing does not by itself make a novel synthetic. The second mistake confines the category to a chat transcript or to a work marketed as autonomous machine creation. Ross Goodwin's *1 the Road* shows why that restriction is too narrow. Its textual event includes sensor data, a custom recurrent neural network, a route, a vehicle, a receipt printer, human design decisions, publication, and translation. Calling it a chatbot novel would erase the apparatus through which the work was made available (Goodwin; CBC Radio). The relevant question is not whether the car or model is an author in a human sense. It is whether these systems participated in conditions that matter to the literary object.

The classification is necessarily conditional. K Allado-McDowell's *Pharmako-AI* visibly stages an iterative exchange with GPT-3, while the complete prompt-output history and exact model configuration are not publicly available. The published edition supports claims about typography, sequence, dialogue, and the creator's attributed description of pruning and

arrangement. It does not support a complete reconstruction of every rejected generation or an account of the model's internal state (Allado-McDowell; Ignota). Jukka Aalho's *Aum Golly* projects make a different configuration visible. Documentation for the third volume includes a corpus record and model-use information that allow unusually specific discussion of generation, selection, revision, translation, and approval. That evidence should not be transferred backward to earlier volumes with thinner archives (Aalho, *Aum Golly 3*; Aalho, "Model Usage"). Stephen Marche's *Death of an Author* presents still another arrangement: human plot and assembly, multiple generative tools, public claims about an AI-written percentage, and no complete public composition archive. The percentage is a documented publisher and creator claim, not an independently audited measure (Pushkin Industries; Harris).

These examples suggest that a literary work is better understood as a *synthetic textual event*. The event is the provisional outcome of relations among human contributors, models, archives, interfaces, institutions, infrastructures, and readers. This formulation does not dissolve the work into its context. It asks which relations are evidenced well enough to change a particular interpretive question. A printed book, a live personalized narrative, and a translated excerpt may all involve generative systems, but they do not present the same object of reading. A fixed edition permits claims about a stable sequence of words and pages. A personalized system may require the critic to specify account tier, date, locale, interface, prompt route, and session conditions before speaking of "the" text.

The concept also clarifies the relation between production and meaning. Barthes's challenge to authorial sovereignty and Foucault's account of the author-function remain central to literary theory because bylines, names, and intentions do not exhaust meaning (Barthes; Foucault). Synthetic Hermeneutics does not restore a sovereign author through technical

documentation. A production record is not an answer key. It can establish that a person prompted, selected, revised, translated, or published. It cannot establish every effect that a reader may find in the finished work. The protocol instead distinguishes genesis claims from interpretive claims: one can be well supported while the other remains contestable.

The Provenance Field and the Five Readings

Genetic criticism provides a useful starting point because it reconstructs composition from surviving witnesses: drafts, proofs, notebooks, correspondence, and versions. Yet machine-learning-mediated texts often lack a stable *avant-texte*. Prompts may be absent, outputs may be deleted, model snapshots may be undisclosed, interfaces may change, and the training archive may be proprietary. A method for these texts must therefore make incompleteness explicit rather than disguising it as an unbroken origin story. The provenance field is the name for that evidentiary landscape.

The field has at least eight dimensions: textual, human, computational, archival, institutional, material, legal, and readerly. The textual dimension concerns the examined language and form. The human dimension records named acts such as prompting, selecting, editing, translating, and publishing. The computational dimension identifies the documented system family, version, interface, settings, retrieval mechanisms, and tools. The archival dimension concerns prompts, outputs, source materials, data claims, and gaps. The institutional dimension includes contracts, publisher statements, prizes, university rules, platform policies, and disclosures. The material dimension includes access, labor, hardware, networks, and resource conditions when they can be linked to the object. The legal dimension distinguishes applicable jurisdiction and legal question. The readerly dimension asks what a particular

audience encountered and knew at the time of reading. No dimension carries automatic priority. The relevant evidence depends on the claim being made.

The field differs from a simple chain of provenance. Chains imply a clean sequence from source to output. Synthetic works more often involve branching, recursive, and contested relations. A creator may prompt a model repeatedly, select outputs, ask another system for variations, edit a passage, and publish an edition whose reader later encounters a changed interface. A publisher's statement may preserve one claim while a creator interview qualifies it. A public archive may verify a model family but not a particular backend snapshot. The task is to record these relations with their warrants, not to force them into one coherent origin narrative.

Synthetic Hermeneutics operationalizes the field through five readings, always taken in this order: textual, production, archive, infrastructure, and responsibility. The order is methodological rather than metaphysical. It begins with the work that readers can inspect and expands the evidence field one step at a time. Each operation has a question, an evidence class, a prohibited inference, and a bounded conclusion.

Textual reading asks what the examined language, form, typography, narrative structure, genre signals, or interface affordances do. It uses the edition, realization, or captured interaction. It may describe repetition, a shift in register, a pattern of address, or a material feature that organizes reading. It must not infer a particular model, prompt, training source, or percentage of machine authorship from those features alone. In *Pharmako-AI*, for instance, the alternation between visually differentiated human and model utterances can support a claim about staged dialogue and formal boundaries. It cannot establish the full order in which every interaction occurred or prove that the two voices possess symmetrical agency.

Production reading asks how a particular textual object became public. Relevant evidence includes dated drafts, interviews, published process notes, edition histories, prompts, output archives, version records, and attributed creator or publisher accounts. The critic records the status of each claim. "Documented" means an inspectable record supports it. "Attributed" means a named person or institution made the claim. "Inferred" identifies a reasoned movement beyond direct description. "Unknown" names a question the available record cannot answer. This vocabulary is more than a stylistic precaution. It prevents a publisher's publicity claim from becoming an independently verified fact and prevents an absent record from becoming evidence of deception.

Archive reading asks what prior materials, cultural traditions, datasets, genres, retrieval systems, and omissions condition the event. It extends intertextual reading without turning resemblance into a source accusation. Longpre and colleagues argue that authenticity, consent, and provenance in AI data practices require serious attention, but a system-level account of web-scale training does not identify whether a named novel was present in a particular model or shaped a particular output (Longpre et al.). Critics must therefore separate training membership, influence, imitation, copying, attribution, and consent. A close parallel may justify comparative reading. It does not independently demonstrate that a work was memorized, scraped, or copied during generation.

Infrastructure reading asks when technical conditions alter form, availability, circulation, reception, or responsibility. The relevant scale may be an interface limit, a model version, an account tier, a moderation rule, a server outage, a translation route, or a platform's personalization system. It need not be an abstract account of "the cloud." Infrastructure studies show that technical arrangements become visible in breakdown, access, and ordinary practice

rather than functioning as a neutral background (Star and Ruhleder; Plantin et al.). The critic should resist converting aggregate estimates of data-center electricity or water use into a per-poem or per-book footprint. Such a calculation requires information about model, workload, facility, allocation, and edition that is usually unavailable (Lei et al.; International Energy Agency).

Responsibility reading asks who knew what, controlled which decision, benefited, could document or correct the process, and could have prevented the relevant harm. It concerns writers, editors, publishers, model providers, platforms, and readers, but does not presume equal responsibility among them. A model may be causally important without being a moral or legal person. This reading brings an ethical question back to the specific issue and actor: a publisher may control disclosure; a provider may control technical documentation; an editor may control the boundary between candidate text and final wording; a platform may control access and moderation. Distributed production should not become distributed evasion.

Together, the five readings make a claim trace possible. A critic can begin with a formal observation, locate any documented production relation, state what archive evidence does and does not show, identify a material condition if it changes the work, and allocate responsibility only after the preceding steps. The trace produces neither certainty nor a score. It produces a record of how far an interpretation may travel.

Evidence Discipline across Style, Translation, Materiality, and Disclosure

The protocol matters most where a plausible observation invites an excessive conclusion. Style is a common example. Readers increasingly describe prose as "AI-like," often referring to smooth transitions, repeated modifiers, generic abstraction, over-complete summaries, or abrupt changes in register. These can be useful textual observations. They can identify a feature worth

comparing across passages, editions, genres, or drafts. They cannot identify a production cause without corroborating evidence. A human writer can produce formulaic prose; a model can produce an unusual or difficult passage; a translator, editor, or interface can produce features that readers mistakenly attribute to generation.

Stylometric and detector research makes this limit especially clear. Zhu and Lei reported more than 80 percent accuracy for Burrows's Delta in a specified collection of human and machine-produced argumentative essays, while also finding that performance changed with text length and feature selection. The result supports a bounded classification task. It does not establish a universal machine signature in novels, translations, lyric sequences, or heavily revised drafts. Liang et al. further show that some GPT detectors disproportionately classified non-native English writing as AI-generated. These findings require a critic or institution to record tool version, date, threshold, corpus, language, genre, text length, preprocessing, error evidence, and appeal route. A score may support an inquiry. It cannot prove fraud, prompt history, machine-written percentage, or the absence of human authorship.

The same distinction applies to intention. Wimsatt and Beardsley's warning against reducing a poem to a recoverable intention remains useful, while current debates add a separate temptation: to treat fluent model output as evidence of belief, understanding, or a stable speaker. Bender et al. caution that fluent language can encourage readers to project grounded knowledge and social understanding onto systems whose conditions do not warrant those projections. Synthetic Hermeneutics therefore separates formal effect, situated address, documented human intention, and model operation. A passage can stage a persuasive voice without proving that a model has an intention. A creator can document a prompt without determining every relation a

reader discovers. The method does not make meaning impossible; it locates different kinds of meaning claim at different evidentiary levels.

Translation makes this separation particularly urgent. Translation was already distributed across source text, translator, editor, publisher, languages, and reader positions before generative systems entered the workflow. A machine-mediated translation adds potential model, interface, prompt, candidate-selection, and disclosure variables. It does not reduce evaluation to the question of whether a target text is fluent. Critics must distinguish source access, semantic adequacy, voice, cultural mediation, production history, and reader response. A monolingual reader may responsibly report the texture and address of a target text, but cannot make the same source-comparison claim available to a bilingual reader. Neither position is uniform, and neither should be treated as a deficit by default.

Recent empirical work supports this caution. Zhang, Xie, and Hu compared selected passages from Chun-chan Yeh's English novel and Chinese self-translation with selected model translations. Their retained native-Chinese participants showed no statistically significant differences in the measured blind-reading outcomes for the study's conditions. A subset of qualitative responses later indicated that hypothetical knowledge of AI involvement could prompt reconsideration. The findings are important, but their boundaries are equally important: the study concerns two selected excerpts, a historical self-translation, particular generation and selection procedures, and a later hypothetical question rather than an actual randomized disclosure intervention. It cannot establish general parity between human and machine translation, professional consensus on fidelity, or a universal disclosure effect. The result instead shows why the critic must keep blind textual response, attributed production information, and measured reception evidence separate.

Disclosure is not an external label added after interpretation. It can become part of the reader's object. Genette's account of paratexts helps explain why a production note, author byline, platform label, or retrospective revelation may shape the threshold through which a work is read (Genette). Yet disclosure should not be treated as a universal correction for opacity. A bare statement that a work was "made with AI" hides more than it reveals. A useful note identifies the system or system family when known, the date or relevant version, the task for which it was used, the human roles that selected or revised output, the edition or release to which the note applies, and material uncertainties. A translation note should additionally identify who compared source and target versions. A live personalized work should identify the relevant session, account, data, moderation, and version conditions.

The timing of disclosure also matters. Prior disclosure can shape expectation before reading. Blind reading can register response before a reader knows the attributed process. Retrospective disclosure can prompt a later reevaluation. Contested disclosure introduces a different problem because the reader may not know which account is reliable. Porter and Machery's experiment on poetry and Raffloer and Green's study of short narratives provide evidence about specific genre, corpus, and disclosure conditions, not a general psychology of AI authorship (Porter and Machery; Raffloer and Green). The methodological lesson is modest: studies should not be combined into a single claim that readers always prefer, reject, or fail to identify AI-generated literature.

Material conditions require equivalent restraint. Models depend on data preparation, annotation, verification, moderation, maintenance, hardware, networks, and energy systems, but it is inaccurate to imagine a single hidden worker or a single planetary cost behind each sentence. Tubaro, Casilli, and Coville identify distinct forms of platform labor that support AI systems,

while environmental scholarship emphasizes the importance of specifying the relevant system boundary. A literary critic may discuss labor or resource use when a record connects those conditions to a named object, interface, or circulation route. The critic should not assign a precise energy, water, or carbon figure to a book merely because it involved a model. In *1 the Road*, the sensor, route, and printer are visible parts of the event. In *Pharmako-AI*, the printed dialogue, the GPT-3 family relation, and editorial selection can be discussed from public records, while facility-level impacts and complete model-call history remain unknown. The unknowns are not an inconvenience to be covered by rhetorical scale. They are part of the interpretation's limit.

Institutions, Authorship, and Bounded Judgment

Institutional disputes often force the questions that criticism can keep distinct. A university may ask whether a student misrepresented a process. A publisher may ask what to disclose in a manuscript or contract. A prize may ask whether a work meets eligibility rules. A platform may triage content using an automated classifier. These are not all authorship questions, even if they are publicly described as such.

Synthetic Hermeneutics separates at least seven maps: causal contribution, literary interpretation, credit, copyrightability, initial ownership, contractual assignment or license, and responsibility. A model's causal contribution does not itself decide literary authorship. A byline does not establish copyrightability. A copyright registration question does not establish ownership of every component of a workflow. A contract does not transform unprotectable material into protected expression. A responsibility judgment does not determine legal liability. Keeping these maps separate does not make institutions indecisive. It prevents them from treating a convenient label as a complete answer.

Current U.S. materials illustrate the value of this distinction. The U.S. Copyright Office's 2025 report treats copyrightability as a case-by-case inquiry into identifiable human expression, selection, arrangement, and modification, while maintaining that purely AI-generated material is not protected under the Office's position (U.S. Copyright Office). The D.C. Circuit's decision in *Thaler v. Perlmutter* concerned an application listing a machine as the sole author and did not decide the threshold of human contribution in mixed workflows (*Thaler v. Perlmutter*). These sources are relevant to U.S. copyrightability and registration practice. They are not universal rules about literary value, credit, moral authorship, contract, or law in other jurisdictions.

The European Union's AI Act makes a different institutional distinction. Article 53 concerns specified obligations of covered general-purpose AI model providers, not a general requirement that every author disclose prompts or that every publisher reveal a full production archive. The European Commission states that the relevant provider obligations entered into application on 2 August 2025, that Commission enforcement powers entered into application on 2 August 2026, and that certain pre-existing models have a transition period until 2 August 2027 (European Commission; European Parliament and Council). The practical implication for criticism is not to convert provider regulation into a generalized book-credit rule. It is to name the jurisdiction, actor, legal question, and date before drawing conclusions.

Where institutions consider sanctions, a minimum evidence record is necessary. It should include the detector or classifier version and date; the validation and local comparison corpus; the threshold; the language and genre; the text length and excerpt boundary; preprocessing; false-positive, false-negative, calibration, and subgroup evidence where available; an appeal route; a decision maker; and corroborating process records. This is not a claim that every allegation will be resolved by documentation. Drafts, logs, testimony, and metadata can also be incomplete or

misleading. It is a claim that automated triage should not become a public fact about authorship without procedural safeguards.

Responsibility belongs at the end of this process because it depends on the object and issue identified earlier. A writer who selected and published output may have duties of accuracy or disclosure. An editor may have duties concerning revision boundaries and reader-facing notes. A publisher may control marketing claims and contractual representations. A provider may control technical documentation, access, and certain design choices. A platform may control moderation, data practices, and visibility. Readers and researchers may have duties of interpretive care, especially when repeating contested claims. These responsibilities are uneven because knowledge, control, benefit, capacity, and preventability are uneven. Distributed causality does not erase accountability.

Discussion and Conclusion

Synthetic Hermeneutics is intentionally limited. It cannot recover deleted prompts, closed training data, undisclosed model parameters, hidden labor chains, or private contracts. It cannot decide whether a work is aesthetically valuable. It does not turn literary interpretation into a vote among evidence layers, and it does not provide legal advice. Its purpose is more practical: to make the critic say which object is being read, which record supports each claim, and where a claim must stop.

The method also faces predictable objections. Some may argue that it is source criticism with extra terminology. The reply is that conventional source criticism does not always account for changing model versions, interfaces, platform policies, personalized outputs, or missing computational archives. Others may worry that production evidence restores authorial intention as the master key. The protocol prevents that outcome by beginning with textual reading and by

marking production claims as documented, attributed, inferred, or unknown. A final objection is that incomplete archives make the method unusable. In fact, incompleteness is the reason for the method. An opacity statement is not a failure to interpret; it is an account of what the available record permits.

The central contribution, then, is not a new vocabulary for machine authorship. It is a disciplined route through a changed literary situation. Critics can read the text closely, reconstruct production only where records warrant it, distinguish archive questions from source accusations, identify material conditions only when they alter the object, and allocate responsibility without treating causal participation as equal moral status. Such a procedure does not make synthetic literature transparent. It makes the limits of transparency legible. That is a necessary condition for criticism that remains both literary and accountable.

Works Cited

Aalho, Jukka. *Aum Golly 3 Poetry Generation Corpus*. 2026. *Zenodo*, https://doi.org/10.5281/zenodo.18205498.

---. "Aum Golly 3: Model Usage." *Aum Golly Documentation*, accessed 17 July 2026, https://aumgolly.com/docs/books/aum-golly-3/model-usage/.

Allado-McDowell, K. *Pharmako-AI*. Introduction by Irenosen Okojie, Ignota Books, 2020.

Barthes, Roland. "The Death of the Author." *Image-Music-Text*, translated by Stephen Heath, Fontana Press, 1977, pp. 142-48.

Bender, Emily M., et al. "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, Association for Computing Machinery, 2021, pp. 610-23. https://doi.org/10.1145/3442188.3445922.

CBC Radio. "An AI and an Artist Go on the Road: 'The Idea Was to Write a Novel with a Car.'" *Spark*, 14 Oct. 2018, https://www.cbc.ca/radio/spark/409-1.4860495/an-ai-and-an-artist-go-on-the-road-the-idea-was-to-write-a-novel-with-a-car-1.4860760.

European Commission. "Guidelines on the Scope of the Obligations for Providers of General-Purpose AI Models under the AI Act." 28 Apr. 2026, https://digital-strategy.ec.europa.eu/en/policies/guidelines-gpai-providers.

European Parliament and Council of the European Union. *Regulation (EU) 2024/1689 of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence*. *Official Journal of the European Union*, 2024, https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

Foucault, Michel. "What Is an Author?" *Language, Counter-Memory, Practice: Selected Essays and Interviews*, edited by Donald F. Bouchard, translated by Donald F. Bouchard and Sherry Simon, Cornell UP, 1977, pp. 113-38.

Genette, Gerard. *Paratexts: Thresholds of Interpretation*. Translated by Jane E. Lewin, Cambridge UP, 1997. https://doi.org/10.1017/CBO9780511549373.

Goodwin, Ross. *1 the Road*. Introduction by Kenric McDowell, translated by Helene Planquelle, Jean Boite Editions, 2018.

Harris, Elizabeth A. "Peering Into the Future of Novels, with Trained Machines Ready." *The New York Times*, 20 Apr. 2023, https://www.nytimes.com/2023/04/20/books/ai-novels-stephen-marche.html.

Ignota. "The Poison Path by K Allado-McDowell." *Library*, accessed 17 July 2026, https://www.library.ignota.org/news/the-poison-path-1/.

International Energy Agency. *Energy and AI*. 2025, https://www.iea.org/reports/energy-and-ai.

Kommers, Cody, et al. "Computational Hermeneutics: Evaluating Generative AI as a Cultural Technology." *Frontiers in Artificial Intelligence*, vol. 9, 2026, article 1753041. https://doi.org/10.3389/frai.2026.1753041.

Lei, Nuoa, et al. "The Water Use of Data Center Workloads: A Review and Assessment of Key Determinants." *Resources, Conservation and Recycling*, vol. 219, 2025, article 108310. https://doi.org/10.1016/j.resconrec.2025.108310.

Liang, Weixin, et al. "GPT Detectors Are Biased against Non-Native English Writers." *Patterns*, vol. 4, no. 7, 2023, article 100779. https://doi.org/10.1016/j.patter.2023.100779.

Longpre, Shayne, et al. "Position: Data Authenticity, Consent, and Provenance for AI Are All Broken: What Will It Take to Fix Them?" *Proceedings of the 41st International Conference on Machine Learning*, vol. 235, 2024, pp. 32711-25, https://proceedings.mlr.press/v235/longpre24b.html.

Plantin, Jean-Christophe, et al. "Infrastructure Studies Meet Platform Studies in the Age of Google and Facebook." *New Media & Society*, vol. 20, no. 1, 2018, pp. 293-310. https://doi.org/10.1177/1461444816661553.

Porter, Brian, and Edouard Machery. "AI-Generated Poetry Is Indistinguishable from Human-Written Poetry and Is Rated More Favorably." *Scientific Reports*, vol. 14, 2024, article 26133. https://doi.org/10.1038/s41598-024-76900-1.

Pushkin Industries. "Pushkin Industries and Stephen Marche Team Up for Groundbreaking AI-Crafted Meta-Mystery." 20 Apr. 2023, https://www.pushkin.fm/news/pushkin-industries-and-stephen-marche-team-up-for-groundbreaking-ai-crafted-meta-mystery.

Raffloer, Gavin, and Melanie C. Green. "Of Love & Lasers: Perceptions of Narratives by AI versus Human Authors." *Computers in Human Behavior: Artificial Humans*, vol. 5, 2025, article 100168. https://doi.org/10.1016/j.chbah.2025.100168.

Star, Susan Leigh, and Karen Ruhleder. "Steps Toward an Ecology of Infrastructure: Design and Access for Large Information Spaces." *Information Systems Research*, vol. 7, no. 1, 1996, pp. 111-34. https://doi.org/10.1287/isre.7.1.111.

Thaler v. Perlmutter. 130 F.4th 1039. United States Court of Appeals, District of Columbia Circuit. 2025. https://media.cadc.uscourts.gov/opinions/docs/2025/03/23-5233.pdf.

Tubaro, Paola, et al. "The Trainer, the Verifier, the Imitator: Three Ways in Which Human Platform Workers Support Artificial Intelligence." *Big Data & Society*, vol. 7, no. 1, 2020. https://doi.org/10.1177/2053951720919776.

United States, Copyright Office. *Copyright and Artificial Intelligence, Part 2: Copyrightability*. 2025, https://www.copyright.gov/ai/Copyright-and-Artificial-Intelligence-Part-2-Copyrightability-Report.pdf.

Zhang, Wenkang, Rui Xie, and Jiaxing Hu. "Reading without Knowing: How Authorship Disclosure Shapes Ethical Engagement with AI-Translated Literature." *Ethics & Behavior*, 6 July 2026. https://doi.org/10.1080/10508422.2026.2696829.

Zhu, Hongao, and Lei Lei. "Detecting Authorship between Generative AI Models and Humans: A Burrows's Delta Approach." *Digital Scholarship in the Humanities*, vol. 40, no. 3, 2025, pp. 1064-75. https://doi.org/10.1093/llc/fqaf048.