Corporate Authorship and Verification Externalities in Automated Mathematical Research

Flyxion
Independent Researcher

November 2026

Abstract

Corporate laboratories now release mathematical manuscripts generated with language model systems in volumes that raise the question whether independent verification can keep pace. This essay separates two questions that public discussion of the November 2026 releases has tended to run together. The first concerns the practical significance of an asymptotic result, here the claim that the matrix multiplication exponent equals 9/4, and is answered by familiar facts about galactic algorithms. The second concerns the institutional arrangement under which claims are produced, verified, and attributed. A simple accounting model shows that when the rate of claim production persistently exceeds the rate of responsible verification, the verification backlog grows without bound. Where verification is undertaken by independent researchers, its costs may be borne disproportionately by parties who neither control production nor capture the principal reputational returns, and the essay states the conditions under which this is an externality and not merely an expenditure. Four verification obligations are distinguished, concerning assertion, encoding, derivation, and semantic correspondence, and it is argued that machine checking of a derivation does not discharge the last. Corporate attribution is shown to be informationally impoverished, in that many distinct contribution histories map to one byline. The central proposition is that ownership of a generative infrastructure does not transfer the epistemic credit later established by collective scholarly labour. The argument is conditional and holds whether or not any particular theorem in the releases proves correct.

Keywords: automated mathematics, corporate authorship, formal verification, verification externalities, matrix multiplication exponent, scholarly attribution

1 Introduction

Mathematical research has always involved three activities that are conceptually distinct even when a single person performs all of them: the formation of conjectures, the construction of proofs, and the verification of those proofs by someone other than their author. A fourth activity, exposition, converts a verified result into something other mathematicians can use. A fifth, attribution, records who is responsible for what. Generative systems reduce the cost of the first two activities substantially, while the remaining three are reduced less, if at all, and the consequences of that asymmetry are the subject of this essay.

The occasion is the public release of mathematical manuscripts by OpenAI. According to the company's community announcement, the public repository contained 722 manuscripts organized into 372 related result families, with Lean formalizations for many but not all of them [3]. Among the results is an October bound showing that the matrix multiplication exponent satisfies $\omega \leq 9/4$ over the complex numbers, which several independent groups have since re-derived or expounded [6, 1]. A further release dated November 24, 2026 claims the matching lower bound for formal arithmetic circuits over every field, and therefore the exact value $\omega = 9/4$ in that model [4]. Earlier, on August 1, 2026, the same organization had released ten results under a corporate byline with no named individuals [5]. The releases are therefore a pattern and not a single event, and an assessment of them should not hinge on the fate of one theorem.

Two kinds of criticism have accompanied the releases. One holds that an improved exponent is practically irrelevant. The other holds that the manner of release imposes costs on the mathematical community and misallocates credit. The first is a conventional observation in theoretical computer science and is largely correct as far as it goes, though it is often stated too broadly. The second concerns a potentially significant change in how mathematical research is organized, and it deserves separate treatment because it survives even if the first is conceded and even if every claimed theorem is eventually confirmed.

2 Practical Irrelevance and Mathematical Relevance

An asymptotic improvement in an exponent does not establish a reduction in running time, memory consumption, energy expenditure, or communication cost at any physically realizable matrix dimension. Consider two algorithms with running times

$$T_1(n) = C_1 n^{2.371}, \qquad T_2(n) = C_2 n^{2.25}. \tag{1}$$

Ignoring lower order terms, the second overtakes the first at

$$n_* = \left(\frac{C_2}{C_1}\right)^{1/0.121}. \tag{2}$$

If the constant factor of the asymptotically faster algorithm is only one thousand times larger, then $n_* = 1000^{1/0.121}$, approximately 6.08×10^{24} in each dimension. A square matrix of that dimension has about 3.7×10^{49} entries, so even one byte per entry would exceed any conceivable storage. This is an illustration and not an estimate of the constants in any actual construction. It shows why a small gain in an exponent can coexist with complete practical infeasibility, which is the standard meaning of the term galactic algorithm. The upper bound is stated as circuits of size $O(n^{9/4+\epsilon})$ for each $\epsilon > 0$, with constants that depend on ϵ [4], and its dependence on ϵ bears on practical feasibility. The strongest formulation is that the asymptotic theorem supplies no evidence of practical superiority at accessible scales, which is weaker than asserting that the construction has been shown to be useless.

Two clarifications keep this observation from being overstated. First, the practical objection and the mathematical claim concern different things. The November manuscript asserts that every formal arithmetic circuit for the product of two $n \times n$ matrices over any field needs at least $\lceil n^{9/4}/2 \rceil$ multiplication or division gates, which together with the known upper bound would settle the exponent in that model [4]. Establishing the exact exponent, if the proof is sound, would be a major mathematical achievement. The stated bound concerns nonscalar multiplication and division gates in a formal circuit model, and a lower bound of this kind says nothing about bit complexity, memory, or hardware. Second, practical irrelevance is not mathematical irrelevance. A reduction of the best known upper bound from the neighbourhood of 2.371 [2] to 2.25 is a substantial change in a quantity studied for more than half a century, and the techniques involved have already prompted elementary re-derivations.

The legitimate complaint is narrower: it concerns the conversion of that achievement into claims about commercial computing capability that the theorem does not establish.

3 Four Verification Obligations

The releases combine natural language manuscripts with machine-checkable artifacts, and the difference between these is easy to lose. It is useful to distinguish four verification obligations. The first is assertion: a mathematical claim is stated in natural language and its proof is checked by readers. The second is encoding: a corresponding theorem statement is written in a proof assistant such as Lean. The third is derivation: a proof of the encoded statement is checked against an inspected set of axioms and dependencies. The fourth is semantic correspondence: an independent examination establishes that the formal statement faithfully represents the advertised mathematical result.

These are related but not strictly sequential. A natural language proof can be verified without ever being formalized, and a formal theorem can be checked correctly while its accompanying exposition remains mathematically inadequate. Machine checking at the third obligation does not discharge the fourth. A kernel that accepts a proof certifies that the encoded statement follows from the encoded hypotheses. It does not certify that the definitions match their intended meaning, that the quantifiers range over the intended domain, or that an apparently general statement has not been specialized in the encoding. Formalization reduces certain verification costs substantially, and the public record already contains independent kernel re-checks of released artifacts [3]. The residual task is to establish semantic correspondence between the formal artifact and the advertised result, a task whose evidentiary standing is not secured merely by increasing the volume of machine-checked proofs. Automated systems may well assist with that task, and the argument does not require that they cannot. It requires only that the resulting judgment be independently warranted.

The scope of the formalization is itself a question about semantic correspondence. The November manuscript states its lower bound for circuits that may use division gates, and the manuscript does not discuss its formalization at all [4]. Whether the accompanying Lean artifacts encode the same class of circuits, or a narrower one

such as division-free circuits, must be read from the repository and confirmed independently, and this essay does not rely on any particular answer. The announcement of the October release states that Lean proofs exist for many but not all manuscripts [3], so a portion of the release rests on the first obligation alone. Independent exposition plays a complementary role. Pass supplies a constructive and elementary proof of the 9/4 upper bound with an accompanying Lean formalization [6], and Amanov, van den Berg, and Zuiddam give an exposition together with a discussion of barriers [1]. These are identifiable scholarly contributions that make an initially machine generated result intelligible, and they exemplify the human work on which the authority of the result will ultimately depend.

4 Verification Demand and Verification Externalities

Suppose a research organization releases manuscripts at rate λ while the relevant scholarly community can responsibly verify them at an aggregate rate μ. Let $B(t)$ be the accumulated backlog of unverified manuscripts. In a simplified continuous model, whenever the backlog is positive,

$$\frac{dB}{dt} = \lambda - \mu. \tag{3}$$

If $\lambda > \mu$ persistently, the backlog grows without bound. With hypothetical rates of $\lambda = 90$ and $\mu = 70$ manuscripts per month and equal verification cost per manuscript, the backlog after one year is 240 manuscripts. These figures are invented for illustration and are not measurements of any organization or review system. What the model establishes is only the growth of the backlog under the stated condition. Whether that condition holds in any real field is an empirical matter, since the actual specialist capacity is unmeasured.

Manuscripts are not interchangeable units, so the model should be generalized. Let manuscript i be released at time t_i and have verification cost c_i measured in specialist researcher-hours, and define the cumulative demand $W(t) = \sum_{i:\, t_i \le t} c_i$. If $v(u)$ is the verification capacity in researcher-hours per unit time, then, assuming an

initially empty backlog and work-conserving verification, the outstanding workload is

$$B(t) = \sup_{0 \le s \le t} \left[W(t) - W(s^{-}) - \int_{s}^{t} v(u)\, du \right], \tag{4}$$

where $W(s^{-})$ denotes the demand strictly before time s. This formulation separates manuscript abundance from verification demand. A hundred short, independently checkable lemmas may impose less work than a single manuscript of 111 pages with several unverified foundational dependencies.

Dependence can be represented by a directed graph $\mathcal{G} = (P, E)$ whose vertices are proof obligations and whose edges record logical dependence. For a claimed theorem T let $\mathrm{Anc}(T)$ be its transitive prerequisites and let S be the set of obligations already verified. The outstanding cost of T is

$$C(T \mid S) = \sum_{p \in (\mathrm{Anc}(T) \cup \{T\}) \setminus S} c(p). \tag{5}$$

The November manuscript is a concrete case. Its proof invokes companion releases of the same date on the geometry of the asymptotic spectrum of tensors and on circuit lower bounds, and it cites an October upper bound and a separate manuscript on the Exponential Time Hypothesis [4]. A reader who wishes to verify the exponent must therefore verify, or trust, a dependency graph and not a single paper. Interdependence cuts both ways. An unverified foundational lemma can obstruct an entire family of results, but verifying it once can also discharge obligations for numerous descendants, so shared dependencies are not intrinsically a burden. The public record also shows that revision is a live phenomenon: secondary reporting describes three manuscripts withdrawn and fourteen revised in place within days of the October release, and the repository maintains revision and citation protocols with earlier versions preserved [3, 7]. The secondary figures are provisional.

Verification expenditure is not by itself an externality. Researchers may examine a proof because doing so advances their own work, offers publication opportunities, or teaches them a valuable technique. Let V_i be the verification expenditure of researcher i, S_i the support received from the releasing institution, and U_i the independent benefit obtained through participation, all in comparable units. The net burden is

$E_i = V_i - S_i - U_i$ and the aggregate uncompensated burden is

$$E_{\text{ext}} = \sum_i E_i. \tag{6}$$

This quantity need not be positive. Verification could yield scholarly benefits exceeding its costs, and an erroneous but highly publicized proof could consume specialist attention while yielding little. The institutional criticism is therefore properly conditional: an externality arises where the releasing organization does not internalize the net costs that its publication decisions impose on others. A separate phenomenon is displacement. Even when examining generated proofs benefits the examiner, the attention spent may displace research programmes that would otherwise have received it, which matters most where the relevant expertise is scarce.

5 Corporate Authorship as an Information-Losing Operation

Corporate authorship does not eliminate responsibility. An incorporated organization can undertake legal obligations, issue corrections, retract publications, and maintain internal accountability procedures. The August release was credited to the organization alone, with a statement that its staff helped prepare the manuscripts and formalize the proofs, that the organization takes responsibility for their correctness, and that the arguments themselves were generated by its system [5]. That acceptance of responsibility is meaningful, though insufficient for the purposes of scholarly attribution.

The more precise objection is that corporate authorship removes publicly accessible distinctions among contributions. Represent the intellectual history of a manuscript as

$$\mathcal{H} = \{(a_i, c_i, r_i)\}_{i=1}^{N}, \tag{7}$$

where a_i identifies a contributor, c_i a contribution, and r_i its role in producing the manuscript. A corporate attribution function $A : \mathcal{H} \to \mathcal{B}$ maps possible contribution histories to institutional bylines. For a given byline b the preimage $A^{-1}(b)$ contains many distinct histories. The byline alone cannot distinguish a manuscript substantially developed by human mathematicians from one generated autonomously and

then edited, nor can it identify who supplied a decisive correction or independently verified a difficult lemma. This is a non-identifiability result for corporate attribution. It does not establish concealment or improper appropriation. It establishes only that the published attribution is insufficient to recover the distribution of intellectual responsibility.

Corporate attribution is not necessarily false attribution, but it is informationally impoverished attribution. It identifies the institution under whose authority a manuscript is released while leaving unresolved the distribution of contributions within it. The distinction matters because scholarly authorship serves not only to allocate prestige but also to preserve the provenance of arguments, to identify relevant expertise, and to establish points of contact for later correction. Institutional responsibility and individual intellectual accountability are complementary and not interchangeable. The point does not require that every mathematician whose published work shaped a model's training receive co-authorship. Disclosure of direct human contributions is a far smaller demand and is sufficient to repair most of the loss.

The credit question can then be put precisely. The relevant issue is not whether an artificial intelligence system deserves authorship, but whether ownership of a generative infrastructure entitles its owner to the reputational returns of a larger epistemic process in which others verify, expound, and extend a claim. Independent expositors may themselves receive substantial scholarly recognition, as the elementary rederivations of the upper bound suggest, so the claim is not that the organization retains all return. It is that the allocation is not determined by provenance.

6 Claim Production and Knowledge Production

Scholarly publication already rewards priority and visibility, and generative systems reduce the cost of an additional manuscript substantially. The marginal cost is not zero, since generation, selection, preparation, and checking all retain costs, but it can fall far below that of human drafting, and an organization can announce many proposed results without waiting for independent confirmation of the preceding ones. Public attention may then track the rate of claim production and not the rate of

justified discovery. Define the rates

$$r_C(t) = \frac{dN_{\text{claimed}}}{dt}, \qquad r_K(t) = \frac{dN_{\text{established}}}{dt}, \tag{8}$$

and, if the terminology of acceleration is wanted, $a_C = dr_C/dt$ and $a_K = dr_K/dt$. There is no necessary proportionality between r_C and r_K.

That excess claim production can impede knowledge production does not follow from the backlog model alone, and a further model states the conditions. Suppose the community has total specialist capacity $H = H_R + H_V$, divided between original research and verification. Let established knowledge production be

$$K = f(H_R) + q\, g(H_V), \tag{9}$$

where f measures the productivity of original research, g the productivity of verification, and q the expected epistemic yield of the claims examined. Since $H_R = H - H_V$,

$$\frac{dK}{dH_V} = -f'(H - H_V) + q\, g'(H_V). \tag{10}$$

Reallocating attention to verification therefore reduces aggregate knowledge production at the margin exactly when $q\, g'(H_V) < f'(H - H_V)$, that is, when the expected yield of the claims examined is lower than that of the research displaced. The condition also admits the opposite case: exceptionally valuable generated mathematics may justify a substantial reallocation of specialist attention.

This does not imply that the matrix multiplication result is false, nor that publishing machine generated research is undesirable. The results have attracted serious independent attention, which is evidence that some of the output is of high quality. The argument identifies an institutional failure mode that can occur even when many results are correct, because the failure lies in the allocation of costs and credit and not in the truth of the statements.

7 What Remains Unmeasured

The factual basis of this essay is deliberately limited. The corporate attribution, the existence of Lean formalizations for many but not all manuscripts, the content of the November manuscript's stated theorem, and the independent expositions of the

upper bound are documented in the sources cited. Several quantities on which the argument ultimately turns are not. The rates λ and μ in any real field, and hence whether claim production persistently exceeds responsible verification, have not been measured. The verification expenditure in researcher hours, the support received from the releasing organization, and the independent benefits to examiners, which together determine E_{ext}, are unknown. The distribution of human contributions to individual manuscripts is not disclosed in the byline, and the practical crossover dimension for any actual construction has not been computed here. These are empirical questions, and the proper response is to measure them, for instance through voluntary reporting of review hours by journals and by groups that undertake re-derivations, and through disclosure by releasing organizations of which humans prepared which manuscripts.

8 Conclusion

Automated mathematical generation alters the relative costs of conjecture formation, proof construction, verification, exposition, and attribution, and it may reduce some of them in the later stages as well, without necessarily altering the institutional mechanisms through which scholarly legitimacy is awarded. The resulting problem is not an abundance of mathematics but a possibly asymmetric distribution of its costs and benefits. A corporation may legitimately own its computational infrastructure, fund a research programme, receive recognition for facilitating a discovery, and accept institutional responsibility for its publications. None of these facts establishes that the corporation is the sole intellectual source of the knowledge subsequently produced through independent verification, reconstruction, and extension. This suggests a principle of non-transferability of epistemic credit through infrastructural ownership: ownership of the productive apparatus does not automatically transfer the authorship of every intellectual contribution mediated by it. The principle holds even if every principal theorem in the releases ultimately proves correct, and it suggests concrete remedies: support for independent verification, named accountability for human preparers, and acknowledgment of the exposition and formalization work that makes a claim usable.

References

[1] A. Amanov, M. van den Berg, and J. Zuiddam. Multiplying Matrices via Polynomials: Exposition and Barriers. arXiv:2610.30054, 2026. https://arxiv.org/abs/2610.30054. Accessed November 25, 2026.

[2] E. Dupont, M. Eisenberger, B. Kozlovskii, A. Mehrabian, F. J. R. Ruiz, A. See, R. Zhou, J. Alman, V. Vassilevska Williams, and M. Balog. Improving the Matrix Multiplication Exponent with Modern Optimization and AlphaEvolve. arXiv:2608.16884, 2026. https://arxiv.org/abs/2608.16884. Accessed November 25, 2026.

[3] OpenAI. First Look at Mathematics Manuscripts from an Internal Frontier Model at OpenAI. OpenAI Developer Community post, October 2026. https://community.openai.com/t/first-look-at-mathematics-manuscripts-from-an-internal-frontier-model-at-openai/1403886. Accessed November 25, 2026. Repository: https://github.com/openai/math.

[4] OpenAI. The Matrix Multiplication Exponent Is Nine Fourths. OpenAI Math Release preprint, November 24, 2026.

[5] OpenAI. Ten Advances in Mathematics and Theoretical Computer Science. Announcement, August 1, 2026.

[6] R. Pass. Matrix Multiplication from Polynomial Packings: A Constructive and Elementary Proof of $\omega \leq 9/4$. arXiv:2610.28627, 2026. https://arxiv.org/abs/2610.28627. Accessed November 25, 2026.

[7] Tech Insider. OpenAI 722 Math Manuscripts from an Unreleased Model. Secondary news report, October 2026. https://tech-insider.org/openai-722-math-manuscripts-unreleased-model-2026/. Accessed November 25, 2026.