



THE BEACON PAPERS • No. 6

Foundational papers for humanity in the AI Age — from Boston

WHO MAY AUTHOR?

*Human Authorship, AI Assistants, and Responsibility
in the Age of Artificial Intelligence*

Judgment, not generation, is the foundation of authorship.

Nguyen Anh Tuan

Co-Founder, Co-Chair, and CEO, Boston Global Forum
Founder and Chief Architect, AIWS

I. Three Theories, One Conclusion, No Agreement

Within a few years of the arrival of capable generative systems, nearly every serious institution in the world wrote a rule about them. The leading scientific journals decided that an AI system may not be listed as an author of a paper. Courts began sanctioning lawyers who filed briefs containing citations that no attorney had bothered to check and no case reporter had ever contained; by the middle of 2026 more than a thousand such orders had been recorded. Universities rewrote their examination codes, some prohibiting these tools outright, others requiring their use. Copyright offices held to the position that protection requires human authorship — a position the Supreme Court of the United States declined to disturb in March 2026 — and then spent the following years working out what that means for a work in which a person and a machine both had a hand.

Read quickly, this looks like consensus. Every institution arrived at the same headline: a machine may not be an author, and responsibility remains with human beings.

What is not generally noticed is that these rules rest on three different and incompatible theories of what authorship is.

The first theory is accountability. The journals reason that a name on a work is a promise that someone will answer for it. On this account, authorship attaches to whoever can be held responsible.

The second theory is professional duty. The courts reason that an advocate must verify what he files, whoever or whatever composed it. On this account, authorship attaches to whoever owed the duty of care.

The third theory is origination. The copyright authorities reason that protection attaches to the human source of the creative expression. On this account, authorship attaches to whoever in fact made the words.

These theories agree on the easy cases and diverge on every hard one. The third is the most consequential in law, and it is the one this paper rejects — though the rejection has to be stated with care, because origin is not irrelevant. It is the wrong kind of origin that the third theory measures.

What matters is not who produced the sentences. It is whose idea the work began as, and whether that person examined what came back, weighed it against his own requirements, and satisfied himself that it set out his thinking fully and accurately. Authorship, on the account offered here, is a fact about whose purpose a work serves and who can be required to answer for it. Where the words first appeared is a question about production.

That account is what this paper sets out, and it is offered against the alternative, not alongside it.

II. The Question Is Not Who Typed

If AI helps produce a speech, a policy paper, a diagnosis, a judgment, or a strategic vision, who is the author?

Put that way, the question sounds like a question about origin — whose fingers, whose first draft, whose sentence came first. That framing is a trap, and it has confused the debate from the beginning. Origin has never been how human societies actually assign authorship in practice. A book taken down from an author's dictation belongs to the author. A statute drafted by legislative counsel belongs to the legislature that enacted it. A judgment written substantially by a clerk is the judge's, and the judge's alone, from the moment it is handed down.

What these cases share is not that the named person produced the words. It is that the named person can be called to answer for them. In every serious institutional setting, authorship is the name we give to the point at which a work becomes answerable — the place where a question about the work can be put to a person, and that person cannot pass it along.

None of this denies that a machine can produce a book. It denies that producing one makes it an author. Authorship in the sense that governs institutions is not a description of who

generated the text. It is a standing to claim a work, to defend it, and to answer for it, and standing is something only a person can hold.

III. Assistance Is Old. This Assistant Is New.

Human civilization has never expected its leaders or its scholars to work alone. Its quiet strength has been the capacity to organize intelligence collectively: advisers, researchers, editors, speechwriters, counsels, and technical experts have shaped the words of nearly every consequential figure in modern history.

It would be convenient to say that no one ever objected to this. It would not be true. When the extent of Theodore Sorensen's contribution to President Kennedy's writing became a matter of public argument, the controversy ran for decades and touched a Pulitzer Prize; Sorensen himself did not give a full account of what he had done until he published his memoir half a century afterward. The dispute was never trivial, and it was never really about who held the pen. It was about whether the work represented Kennedy's own thought and judgment, or was a reputation assembled on his behalf by someone else.

That is precisely the question now, at civilizational scale.

But the analogy to human assistants has to be handled honestly, because at one point it breaks. A human assistant is himself a responsible person. He can be named, questioned, dismissed, sued, subpoenaed, and shamed. Responsibility in the old arrangement was concentrated in the principal, but it was never sealed there; some portion of it could always be traced outward to another human being who also had to live with the work.

An AI assistant cannot be reached in any of these ways. It has no reputation to lose, no career to end, no conscience to trouble. Nothing can be recovered from it afterward.

This is usually presented as an objection to the continuity of the tradition. It is the reverse. Because nothing can be traced outward, responsibility does not thin as the assistance grows — it concentrates. The more capable the assistant, the more completely the moral weight of the finished work rests on the single human being who approved it.

IV. What Kind of Assistant This Is

An AI assistant works at any hour. It reads more in a minute than a research team reads in a week, holds an argument steady across a long document, and will produce a fifth alternative as readily as a first. Within a few years a capability of this kind will be unremarkable, available to a prime minister and to a student, to a hospital director and to a village official.

It also asserts things that are not true, and it asserts them in the same confident register it uses for things that are. A system built to produce plausible continuations will produce a plausible citation, a plausible statistic, a plausible precedent, and it carries no internal signal

separating what was invented from what was remembered. Research is reducing how often this happens and will keep reducing it. No responsible institution should assume that rising capability alone will end it.

The argument here, however, does not depend on how often it happens. A system wrong once in a thousand times still gives the reader no way to know which sentence is the one. Accuracy that cannot be told apart from error, from the outside, still requires someone on the inside who checked.

Most discussion of AI and authorship leaves this out, and the omission is costly. Someone has to stand between the fluency of the machine and the trust of the public, and that someone can only be a person. The requirement does not relax as the systems improve. It is what makes their improvement usable.

V. The Conditions of Human Authorship

Using AI does not diminish human authorship. Authorship is preserved when a person sets the purpose, supplies the values, exercises independent judgment on what the assistant offers, rejects as readily as accepts, verifies what the work asserts as fact, approves it in final form, and accepts responsibility without reservation.

None of these are acts of production. All of them are acts of judgment. A person who does these things is the author of the work in the sense that matters morally and before the public, regardless of how many of its sentences first appeared on a screen rather than in a mind. He should be recognized as its responsible human author in law as well, even where particular legal regimes still apply a narrower test.

Judgment, not generation, is the foundation of authorship.

VI. When Authorship Is Lost

A principle that only permits is not a principle. The conditions set out above can fail, and when they fail the human name on the work becomes a false claim.

Authorship is lost when a person approves what he has not read. It is lost when he lets assertions of fact stand that he has not verified and could not verify if asked. It is lost when he cannot explain the reasoning of the work, defend its conclusions under challenge, or answer for the choices it makes. And it is lost when he signs in a setting that requires assistance to be disclosed and does not disclose it.

The last of these deserves emphasis, because the temptation is greatest there and the harm least visible. A name on a work is a representation made to whoever receives it. When the representation is false, the injury is not to the machine, and not to the author. It falls on the reader who relied on it.

There is a second injury, and it falls on the person who signed. Leadership erodes when responsibility is delegated together with judgment. A leader who signs words he neither understands nor could defend has surrendered the very authority his signature was meant to carry, and he will usually not notice, because nothing visible changes on the day it happens.

There is a plain test, and anyone who uses these instruments should be able to pass it on any given day. *Could you defend this work, line by line, under hostile and informed questioning, without once saying that the assistant produced it?* If the answer is yes, you are the author. If the answer is no, you have published something you do not actually stand behind, and the signature is a courtesy you have extended to yourself.

VII. Whom This Protects

Rules of authorship have never existed for the benefit of authors. They exist for the people on the other side of the page.

The citizen reading a national policy is entitled to know that a responsible official examined it. The court receiving a filing is entitled to assume that the cases cited in it exist. The examiner assessing a dissertation must be able to conclude something about the candidate's mind. The patient receiving a diagnosis must be able to hold a physician to it. In every one of these settings, the value of the human signature lies entirely in what it tells the recipient about the care that was taken.

It follows that the duty of disclosure is not uniform, and any workable rule has to recognize a gradient. A political speech has always been understood to be a collaborative product, and no one is deceived by it. A scientific paper is a claim about what its authors did and found, and assistance that touches the finding must be declared. A judicial opinion or a legislative text carries the authority of an office, and the public interest in knowing how it was produced stands at its highest. A student's examination is a measurement of a single mind, and there disclosure is not merely required — it is the whole substance of the exercise.

The question in each setting is the same: what did the recipient reasonably understand the signature to mean?

A clear rule protects honest innovation as well. Where authorship is left undefined, the asymmetry runs the wrong way: careful people hesitate to use these instruments at all, while careless people shelter behind them. Definition reverses this. It lets the responsible work openly, and leaves the irresponsible without cover.

VIII. The Principle

For the whole of recorded history, assistance of the highest quality has belonged to a few — heads of state, heads of institutions, the wealthiest universities. That is now ending. A student, a physician, a mayor, an entrepreneur, and an ordinary citizen may each hold an

assistant of a capability that a head of government could not have commanded a decade ago. Governed well, artificial intelligence makes exceptional assistance universal rather than exclusive, and this is among the most consequential goods it offers. No institution should be permitted to withdraw it from the people in its care for reasons of its own convenience.

The Constitution for Humanity in the Age of Artificial Intelligence should state the matter as a right, joined to a duty, and bounded by a limit.

Every human being has the fundamental right to employ artificial intelligence as an assistant in thinking, learning, writing, creating, governing, and serving society.

No government, employer, university, or institution may prohibit the responsible use of AI assistance merely because it changes established methods of work. Institutions should regulate responsibility, not suppress assistance — except where the unaided work of a single mind is itself the thing being measured or attested, in which case the restriction must be declared openly and confined to that purpose.

Human authorship is preserved wherever a human being exercises independent judgment over the work, verifies what it asserts, approves it in final form, and accepts full responsibility for its contents.

Human authorship is forfeited wherever a person claims a work he cannot explain, defend, or answer for, or conceals assistance in a setting that requires its disclosure.

Artificial intelligence may assist in any of this. It may not author, and it may not be held to account. Accountability belongs to human beings, and may not be transferred, delegated, or shared with a machine.

Stated in this form, the principle protects innovation and human dignity at the same time. It asks no one to use less AI. It asks everyone who uses it to remain answerable for what they sign — which is the oldest requirement of public life, restated for new instruments.

IX. The Capacity to Judge

Everything above rests on a single assumption: that authorship survives the use of AI because human judgment survives it. But judgment is not a possession. It is a capacity, and capacities are maintained by exercise.

Modern constitutions are written almost entirely in the language of rights. This is their strength and also their blind spot. A right describes what a person may do. It says nothing about whether the person still holds the capacity that makes the right worth having. A right to use an AI assistant, protected on its own, guarantees nothing about the judgment that makes authorship real. A society that protected the first and neglected the second would discover, within a generation, that it had preserved the form of human authorship and lost its substance.

A constitution for the Age of Artificial Intelligence must therefore do something constitutions have not usually done. It must identify not only the rights human beings hold, but the capacities human beings are responsible for keeping.

The obvious objection is that this worry is ancient and has always been wrong. Writing was said to destroy memory. The calculator was said to destroy arithmetic. Neither loss proved worth mourning. The objection would hold if judgment were a skill of the same kind. It is not. Losing the ability to calculate in one's head costs a person a convenience. Losing the capacity to judge costs a person the standing to be held responsible at all — and responsibility is not one skill among others. It is what makes someone a person before the law, before an institution, and before another human being. A civilization can afford to forget arithmetic. It cannot afford to produce leaders, physicians, judges, and citizens who can no longer be answerable, because there would then be no one left of whom to ask anything.

The right to use artificial intelligence therefore carries a duty running alongside it: to preserve the capacity that makes authorship real. That duty falls on individuals, but individuals cannot discharge it alone. Education determines whether judgment is ever developed. Institutions determine whether it is ever exercised. Leaders determine whether the people around them are permitted to think, or only to approve.

How this duty is to be met — in schools, in governments, in professions — is a question large enough to require its own treatment. The Beacon Papers will take it up next. What must be settled here is that the duty exists, and that it is not optional.

X. The Human Signature

The future will not belong to those who refuse these instruments, nor to those who hand their judgment over to them. But that is a familiar thing to say, and it is addressed to leaders, who are not the ones at risk.

The people at risk are on the receiving end — the reader, the citizen, the litigant, the patient, the examiner — who cannot see how a document was made and have no choice but to trust the name at the bottom of it. Everything in the Age of Artificial Intelligence that makes writing faster, cheaper, and more abundant makes that name more important rather than less. It is now the only part of a document that a machine cannot supply.

A signature has always been a promise: *I have considered this, and I will answer for it*. Better assistants do not make that promise harder to keep. They only make it easier to break.

The measure of civilization in the Age of Artificial Intelligence will not be the intelligence of its machines, but the judgment of its people.

BEACON PRINCIPLE No. 6

AI may assist.
Only humans may author.
Only humans may be accountable.

SOURCES

Section I — journal authorship policies. Committee on Publication Ethics, *Authorship and AI Tools* (position statement, 2023); International Committee of Medical Journal Editors, *Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals*. Both hold that AI tools may not be listed as authors because they cannot take responsibility for the work; COPE adds that, not being legal entities, such tools can neither declare conflicts of interest nor hold or assign copyright. The same position is taken by WAME, EASE, the Council of Science Editors, and the major scholarly publishers.

Section I — judicial sanctions. Sanctions have been imposed under Federal Rule of Civil Procedure 11, Federal Rule of Appellate Procedure 38, and the courts' inherent authority. See *Coomer v. Lindell* (D. Colo. 2025); *Fletcher v. Experian Information Solutions* (5th Cir.); *In re Nwaubani* (4th Cir.). A public database tracking such orders had recorded more than 1,100 instances by mid-2026. Courts have held that they need not find that AI was used before sanctioning a filing containing fabricated authority: the duty of candor and the obligation to verify are longstanding and independent of the instrument employed.

Section I — copyright authorities. United States Copyright Office, *Copyright and Artificial Intelligence, Part 2: Copyrightability* (January 2025); *Thaler v. Perlmutter*, 130 F.4th 1039 (D.C. Cir. 2025), holding that the Copyright Act requires every eligible work to be authored in the first instance by a human being; certiorari denied, March 2026. The Office's position is that prompts function as instructions conveying unprotectable ideas, and that protection turns on human creative control over the expressive elements of the work. This is an origination test, not an accountability test, and it is the position this paper contests.

Section III — Sorensen and Kennedy. *Profiles in Courage* (1956) was awarded the Pulitzer Prize for Biography in 1957. In December 1957, on ABC, the columnist Drew Pearson stated that the book had been ghostwritten; Kennedy's representatives demanded a retraction, and Sorensen swore an affidavit limiting his role to the assembly and preparation of research materials. In *Counselor: A Life at the Edge of History* (2008), Sorensen acknowledged that he had prepared first drafts of most chapters and helped choose the words of many of the book's sentences.

A NOTE ON THE MAKING OF THIS PAPER

This paper was prepared with the assistance of the AI systems of AIWS Lumina Lab, under the conditions set out in Section V. The ideas, the purpose, the detailed requirements, the argument, and the positions it takes are the author's, given to the assistants as a leader gives instructions to those who work for him — the arrangement Section III describes and Section V recognizes. Drafts were rejected as readily as they were accepted; the author set the revisions required and accepted the work only when it met them. The sources listed above were checked against the primary record rather than accepted as offered. The author has read every sentence that appears here, and will answer for each of them.

This note is not a formality. It is the standard the paper asks of everyone else, applied first to itself.

Boston, August 6, 2026