

A Declaration for Human Memory in the Age of Artificial Intelligence

AI stands on the shoulders of giants — trained on the labor, insight, and imagination of countless human minds.

Executive Summary: A Declaration for Human Memory in the Age of AI

The Problem: A Future Without Names

Generative AI is reshaping creative industries by producing content — images, text, music, and more — based on vast datasets of human-made work. Yet, it often fails to recognize or reward the original creators. This omission is not just an oversight; it constitutes a systemic erasure of human contribution, threatening cultural memory, creative dignity, and the social bonds of authorship.

The Proposal: Attribution as Ethical Infrastructure

This declaration advocates for a foundational shift: embedding **attribution** — the crediting of human creators — into AI systems. Attribution is not merely a courtesy; it is a legal, moral, psychological, and cultural necessity that:

- Upholds **moral rights** recognized in international law (e.g., the Berne Convention).
- Supports **personal dignity**, identity, and mental well-being (backed by neuroscience and social theory).
- Prevents **epistemic and cultural erasure**.
- Serves as the **first layer** for enabling fair economic models (e.g., micro-royalties).

Attribution-Driven Compensation: Micro-Royalties

The declaration proposes a scalable system of **micro-royalties** — small payments based on AI output influence. This is enabled by:

- New attribution-tracing techniques (e.g., influence functions, Shapley values).
- Logging outputs with reference to their most influential sources.
- Distributing royalties via collective management platforms.

The rollout is phased:

1. **Pilot programs** with opt-in licensing and nominal payments.
2. **Standardization and regulation** of attribution features.
3. **Global scale-up**, possibly including AI training registries and universal attribution protocols.

Why Attribution Matters Beyond UBI or Robot Taxes

While Universal Basic Income (UBI) and robot taxes address **economic displacement**, attribution addresses:

- **Recognition** and creative identity.
- **Agency**, by keeping creators active participants in cultural production.
- **Moral rights**, which money alone cannot replace.
- **Social cohesion**, by preserving the human face of creativity in an AI-driven world.

Attribution and economic redistribution are **complementary**, not redundant: one feeds the body, the other feeds the soul.

Political and Social Viability

The attribution framework appeals across ideological lines:

- **Progressives** see justice for exploited labor.
- **Conservatives** see protection of property and incentives.
- **Centrists** see a balanced engine of innovation and fairness.

It is not a partisan agenda — it is a **foundational ethic** for the AI age.

Call to Action

The declaration invites:

- **Developers** to build attribution systems into AI.
- **Policymakers** to legislate source transparency and micro-licensing.
- **Creators and publishers** to organize, register, and demand rights.
- **Citizens** to favor and demand attribution-aware technologies.

Attribution can transform AI from a force of dispossession into a force of collaboration. It ensures that as machines learn and create, they do so in partnership with — not in place of — the human minds that shaped them.

Introduction: When Humanity is Forgotten

Imagine a world of abundance. Machines write, paint, code, diagnose, invent, and strategize — faster and cheaper than any human. Basic income ensures a material well-being. But something deeper begins to vanish: **the sense of matter**. The joy of discovery. The act of contributing and creating. Culture expands, but without roots. Knowledge grows, but without names. We are surrounded by ideas — yet authors disappear.

This is not a distant dystopia. It is the world we are already building.

Generative Artificial Intelligence (AI) is transforming all industries, art, literature, code, music, etc. These systems leverage a vast corpora of human-created works to produce novel outputs, often without credit or reward to the original creators (academic.oup.com). This

dynamic has sparked lawsuits and ethical outcry, as creators fear a “theft of creative and intellectual authorship” by AI developers (publishers.org). AI generates — but does not remember. It responds — but obscures its sources. This is not progress. This is **engineered amnesia**.

Even if universal basic income secures our bodies, it cannot sustain our spirits. **Creativity is not a luxury. It is the human capacity to shape, to mean, to matter.** To be named — to be recognized — is to remain part of history.

We do not propose slowing AI down.

We propose making it more responsible, more interpretable, and more human-aligned by enabling **Attribution**. Attribution paves the way for collaboration — not conflict — between creators and developers. By enabling ethical access to copyrighted work, **it can accelerate innovation, not hinder it.** It turns resistance into partnership, opacity into traceability, and simulation into shared progress.

Let us build systems that remember who we are, not just what we’ve made.

At stake are fundamental human values: the dignity of creative labor, the moral rights of authors, and the social fabric of recognition and reward that underpins cultural production and human well-being. This declaration argues that **attribution** – the act of identifying and crediting human creators when their work informs AI-generated content – is a foundational (though partial) solution to these challenges. By embedding robust attribution into AI systems, we reaffirm human **agency** and **creativity**, uphold the **moral rights** of creators, and lay the groundwork for fair **compensation** (e.g. micro-royalties) when AI exploits human creative contributions.

In what follows, we present the scholarly and philosophical rationale for an attribution-centric approach, outline a pragmatic framework for attribution-based micro-royalties, and position attribution as complementary to broader economic measures like universal basic income (UBI) or “robot taxes.” We also address potential flaws and counterarguments, refining the declaration’s vision with academic rigor and moral clarity.

This is not a closed doctrine — it is an **open declaration**.

A beginning, not an end. We invite artists, developers, policymakers, lawyers, scholars, philosophers, engineers, and citizens to become its co-authors.

Together, let us shape a future where AI amplifies human creativity — without erasing the humans behind it.

Human Creativity, Dignity, and Moral Rights

Human creativity is not only an economic activity but a deeply personal and cognitive one, tied to our sense of identity and dignity. From a **philosophical** perspective, acts of creation are an expression of the self – a view echoed in personhood theories of intellectual property which hold that an author’s work embodies their personality and will (copyright.gov).

International human-rights law explicitly links creative labor to human rights and dignity. Article 27 of the *Universal Declaration of Human Rights* declares that “everyone has the right to the protection of the moral and material interests resulting from any scientific, literary or artistic production of which he is the author” ([humanrights.com](https://www.humanrights.com)).

In the context of AI, stripping away an author’s attribution treats the creator as a mere means to train algorithms, rather than as an end worthy of recognition. This encompasses both **moral rights** (reputation, attribution) and **material rights** (economic interests). Attribution is thus not a trivial courtesy – it is tied to the author’s *moral interests* and dignitary rights on a global legal stage. Many legal systems recognize a **moral right of attribution** (sometimes called the *right of paternity*), whereby creators have an inalienable right to claim authorship of their work.

In the Berne Convention (Art. 6bis), for example, authors hold “the right to claim authorship of the work” even after economic rights are transferred ([wipo.int](https://www.wipo.int)). Such provisions stem from the belief that acknowledging the creator is a matter of justice and respect. Legal scholars note that moral rights “*affirm the basic human dignity of authors and artists everywhere*” ([researchgate.net](https://www.researchgate.net)), recognizing that a creative work is an extension of the person who created it. Failing to credit a human for their creative output not only violates these moral rights but can be seen as an affront to human dignity – akin to plagiarism on a civilizational scale.

The principle that people deserve the fruits of their labor also has deep philosophical roots. John Locke’s labor theory of property holds that individuals are entitled to ownership of the product of their labor ([cdn.mises.org](https://www.cdn.mises.org)). Applied to creativity, this suggests that when an artist’s work is used (even indirectly, as training data for AI), they maintain a natural claim to that contribution. Modern interpretations of Locke extend this to intellectual property, viewing creative works as the “fruits” of one’s mental labor. While Locke’s theory traditionally justifies *property*, its spirit also supports **attribution**: at minimum, a creator has a right to be identified as the source of their labor’s fruits. In essence, both **personality-based** and **labor-based** theories of rights converge on a common insight: respecting creators – by crediting their authorship – is an ethical imperative grounded in human dignity and justified claim.

From the vantage of **cognitive neuroscience and psychology**, the importance of recognition is further underscored. The human brain responds powerfully to social reward and acknowledgment. Neuroimaging studies have found that “*the ventral striatum is activated by social rewards such as praise*” ([pmc.ncbi.nlm.nih.gov](https://pubmed.ncbi.nlm.nih.gov)) – in other words, being recognized and appreciated triggers the brain’s reward circuitry similarly to monetary rewards. This neural evidence supports what artists and authors have always reported anecdotally: recognition matters. It satisfies a fundamental psychological need. In Maslow’s hierarchy of needs, once basic safety and livelihood are secured, humans crave **esteem** – which includes status, fame, recognition, prestige, and attention ([simplypsychology.org](https://www.simplypsychology.org)).

Creative people often derive a sense of purpose and identity from the acknowledgment of their work. The **theory of recognition** in social philosophy (espoused by Axel Honneth and others) likewise argues that human dignity emerges through our relationships and our recognition as meaningful participants in society ([psc](https://www.psc)).

To deny someone appropriate credit for their creative contribution is to deny them a portion of this social esteem, potentially undermining their sense of self-worth and belonging in the creative community. Thus, attribution is not only a legal formality; it is tied to humans' **mental well-being and motivation**. It nurtures the intrinsic reward that fuels further creativity. Indeed, fields like academia have long functioned on a "reputation economy" where citation and attribution are the coin of the realm – a key incentive for authorship. If generative AI systems bypass attribution, they threaten to "*break the reputation ecosystem*" that motivates creative and scholarly endeavors (scholarlykitchen.sspnet.org).

In summary, a robust commitment to attribution upholds *human dignity, agency, and moral rights*. It acknowledges that behind every AI-generated image, paragraph, or melody reminiscent of some style lies human creators who contributed to that style or corpus. Attribution treats that creator as a person with rights and honor, not as an exploitable data point. It operationalizes a core ethical principle: **justice** requires giving each their due – *suum cuique*. In this case, what is due is recognition.

Attribution as a Foundation for Ethical AI

While attribution alone cannot resolve all disputes over AI, it forms the ethical foundation for addressing the social challenges posed by generative AI. **Why focus on attribution?** Because it directly addresses issues of **transparency, accountability, and respect** in the AI-human relationship. Attribution answers the critical question: "*Who (or what) was the source of this content?*" In doing so, it confers *meaning* and *legitimacy* to human creative work in the age of AI.

First, attribution provides **transparency and truthfulness** in AI outputs. Today's generative models often present information in an authoritative tone with no indication of sources (academic.oup.com). Users may be misled to believe the AI "invented" text or images out of thin air, obscuring the human knowledge it was built on. For example, when ChatGPT was asked a question drawn from a famous law review article, it produced an answer with the key arguments but "*gives attribution to neither the authors' [names] nor the paper's title*" (academic.oup.com). This lack of provenance not only shortchanges the authors, it also robs the user of context – they cannot follow up on the original source or judge the credibility of the information. By implementing attribution (e.g. citations or at least listing influential sources), AI systems become more honest about their debts to human creators. This fosters **trust**: users can see *why* an output is authoritative (it draws from recognized work) and creators see that AI is not a plagiarist but an amplifier that still points back to them. In scholarly AI applications, the inability to reference sources is seen as a "significant challenge"; conversely, if AI tools can "*reference back to the source material in their outputs*", it preserves the "reputation economy" and accuracy of knowledge dissemination (scholarlykitchen.sspnet.org).

Second, attribution helps restore **human agency and credit** in the creative process. One societal fear is that as AI generates content, human creators become invisible or devalued. But attribution ensures that whenever AI output significantly imitates or builds upon a particular artist, document or dataset, the humans behind those influences are visibly credited. This keeps humans "in the loop" of creativity, not as passive bystanders but as

recognized co-creators of the cultural output. It is a way of saying: *AI did not emerge ex nihilo; it stands on the shoulders of human creativity*. Such recognition can have concrete effects on agency. For instance, if artists know they will be credited (and potentially compensated) when AI uses their style, they regain a measure of **control**: they can decide whether to permit such use and feel secure that their legacy and name will persist in the AI-generated derivatives. Philosophers of law often speak of **desert** – the principle that rewards (or recognition) should flow to those who deserve them by virtue of contribution. Attribution operationalizes desert in the AI context by linking output back to contribution. It combats the alienation creators might feel when their work is absorbed into a faceless AI model. Instead of disappearing into a black box, the creator's identity *re-emerges* alongside new content. This also upholds **integrity** and truth in authorship: it prevents misrepresentation of AI-generated work as purely machine-made when in fact it is a mosaic of human creativity.

Third, attribution is an **ethically minimalist** and widely agreeable principle, which makes it a promising focal point for consensus. Regardless of one's stance on stronger interventions (such as strict copyright enforcement or bans on AI training data usage), almost all stakeholders can agree that giving credit is fair. Attribution does not inherently impede innovation – it asks for **recognition, not restriction**. In fact, even open-access and Creative Commons communities, which favor free sharing, maintain that author attribution is essential (e.g. the CC BY license requires credit to the author as a condition of reuse) (scholarlykitchen.sspnet.org). Thus attribution can be seen as a common-ground ethical norm, a baseline expectation for **responsible AI** similar to how we expect academic writing to credit sources. By enshrining attribution as a norm in AI, we create a culture of **reciprocity**: AI benefits from human culture, and in return it elevates and acknowledges the humans it learns from. This cultural norm could help ease the current adversarial tension between creators and AI firms.

Crucially, we position attribution as **foundational but partial**. It is *foundational* because without it, other solutions (like payments to creators) lack a principled basis and mechanism. **Attribution is the prerequisite for compensation**: one cannot remunerate thousands of contributors without knowing who they are and how they contributed. It is also foundational for preserving human dignity in an automated loop; even if society implements safety nets like UBI, failing to credit human creators would leave a moral wound – the sense that human creativity has been subsumed without acknowledgment. At the same time, attribution is *partial*: it alone doesn't pay the bills or ensure livelihoods for creators impacted by AI. A starving artist cannot live on credit alone (literally or figuratively). Therefore, attribution should be seen as the first step of a broader framework. It **complements** other measures (legal, economic, technical) by focusing on the layer of meaning, recognition, and moral agency. In ethical terms, attribution addresses what philosopher Joseph Raz might call the *expressive harm* of AI appropriation (the insult to human dignity and identity), while economic measures address the *material harm*. Both need attention. This declaration asserts that any comprehensive policy for generative AI must include attribution at its core, lest we solve the monetary issues but still erode the soul of human creativity.

In sum, attribution stands as a mediating principle between humans and AI – a way to ensure that humans are *seen* and *respected* by the very technologies that emulate them. It is a modest ask with profound implications: by attributing, we keep the human author in the

narrative of creation, preserving a space for **human dignity, creative agency, and truth** in how AI-generated content comes to be.

Attribution-Driven Compensation: The Micro-Royalty Framework

Recognizing creators in AI outputs is a moral imperative, but it also opens the door to **practical compensation mechanisms**. This section proposes a scalable framework for **attribution-based micro-royalties** – small payments to creators when their works are used in AI training or appear in AI outputs – and explains how this model can be implemented fairly and gain political support. While ambitious, this approach is designed to start modestly and **evolve over time**, aligning incentives across stakeholders.

3.1. Why Micro-Royalties as a Starting Point?

Micro-royalties (fractions of a cent or other small units paid per use) are a logical starting point because they mirror how value is generated in the AI context – in *high volume and diffusely*. Just as streaming music platforms pay artists tiny per-play fees that accumulate with scale, generative AI could pay creators each time their work significantly influences an output. No single AI-generated image or paragraph is likely to earn a creator more than pennies, but across millions of generations the revenue could be meaningful. Starting with micro-amounts also lowers the barrier to adoption. It avoids the pitfalls of large, one-time licensing fees which AI developers may resist as too costly or rigid. Instead, it says: “Train on anything, but when you *use* it, pay a little back.” This model aligns with the fact that AI outputs have **derivative value** – value that owes to underlying creative works. “*Flattery won’t pay the bills*”, as one artist quipped when AI mimicked his style without compensation (goizueta.emory.edu). Micro-royalties turn flattery into a tangible (if small) reward. Crucially, recent research indicates public support for such a pay-per-use approach. In consumer experiments, people “*were willing to pay more for a product if they knew the artist would get compensated*”, and they preferred a usage-based royalty model over a flat fee for training data (goizueta.emory.edu). This suggests that micro-royalties not only make economic sense but also resonate with **intuitive fairness** – users and creators alike see it as more just that payment correlates with actual AI usage of an creator or author’s work (goizueta.emory.edu). In short, micro-royalties are **scalable** (they can handle millions of micropayments), **flexible** (payout is proportional to usage), and **perceived as fair** (goizueta.emory.edu). They set the stage for a compensation system that grows with the technology.

3.2. Attribution Data as the Backbone of Remuneration

How can we practically implement micro-royalties? The answer lies in **attribution data**. If AI systems are designed (or retrofitted) to track which inputs significantly contributed to a given output, that metadata can feed directly into a royalty distribution system. Imagine a generative model that, alongside its output, produces a list of the top N sources (artists, authors, datasets) that influenced the result – essentially an AI-generated citation list. For example, a text model might say an answer was 5% inspired by Author A’s writings, 3% by Author B, etc., based on similarity or prompt usage (academic.oup.com). An image model

might identify which training images its output most closely resembles. Such attribution data could be computed using techniques like *embedding similarity*, *Shapley values*, or other game-theoretic contribution metrics (arxiv.org). In fact, researchers have begun developing methods to “*quantitatively determine contributions*” of training data to outputs. One recent proposal uses cooperative game theory to allocate credit among training images and then split revenue accordingly (arxiv.org). While technical challenges remain (AI companies note it is non-trivial to trace outputs to specific training data (academic.oup.com), the path is clear: better attribution logging and model interpretability will make it feasible to *identify* who should be paid.

Once attribution for an output is determined, a **royalty engine** (perhaps administered by a collective management organization or an open platform) could micro-charge the user or AI service and split the fee among the identified contributors. For instance, if an AI-generated image is used commercially (say printed on a poster for sale), a tiny royalty might be owed to each artist whose style influenced it. If a passage of AI-written text clearly derives from a few public-domain novels plus one copyrighted book, the system could allocate a portion of earnings to the latter’s author or publisher. Attribution thus becomes the **lookup key** for remuneration: it links output back to human authors, enabling automatic or semi-automatic payment flows. Importantly, attribution data also enables **auditing and accountability**. Creators could see when and how frequently their works are used by AI, bringing much-needed transparency. It’s worth noting that even if full granularity is hard, approximate credit is still valuable – e.g. just naming the top source authors for each output would allow a simple split of a royalty pool. This is analogous to how performance royalties in music are distributed based on setlists or radio play logs: not perfectly precise but sufficient for fairness over the long run.

3.3. Fair Prioritization: Who Benefits First?

A pragmatic rollout of attribution-based compensation should prioritize those creators most in need and most clearly affected. **Professional creators** – people who make a living (or attempt to) from their art, writing, music, photography, etc. – are a sensible first tier. These individuals have both the greatest stake and the easiest-to-identify works. Many have copyrighted portfolios, publisher relationships, or are members of registries/collectives, making it straightforward to route payments to them. They are also the ones whose livelihoods are at risk from unbridled AI generation. By contrast, the internet is also full of casual or anonymous content (forum posts, me mes, etc.) that AI might ingest; trying to immediately compensate every bit of that would be intractable. Prioritization avoids “analysis paralysis” by focusing on the **high-impact cases**: if an AI model is heavily trained on, say, the artworks of Greg Rutkowski (a digital artist) or on New York Times articles, then Rutkowski and the journalists are first in line to get attribution and royalties (goizueta.emory.edu). This is both fair and politically prudent. It shows goodwill toward creators who have vocally objected to AI’s use of their work, thereby addressing the most visible injustices and building support for the system.

Concretely, the initial implementation could involve **opt-in licensing schemes**: creators (or their publishers) register works with an AI platform or industry body, agreeing to allow AI training/use in exchange for attribution and micro-payments. This would quickly create a pool of professional-grade content that is *licensed and compensated*. AI companies might even

prefer using that pool, to avoid legal uncertainty from unlicensed data. Over time, as the system scales and technology improves, it could expand to cover more content and even automatically recognize unregistered contributions (similar to how YouTube's ContentID pays music rights-holders when their music is detected in videos). Throughout, fairness must be monitored: mechanisms should prevent an outcome where only the top 1% of famous creators get all the benefit. One approach is a **royalty floor or bonus** for smaller creators – e.g. a minimum payout if your work is used at all, or extra weight given to emerging artists to ensure they are not drowned out by only the most “popular” influences. The data from attribution logs can inform these policies, as we see which creators contribute widely to AI outputs. The overarching principle is **inclusivity with priority**: start with the creators who contribute the most and depend on their creative income, then broaden the circle.

3.4. Appealing Across the Political Spectrum

Any sustainable AI policy must be **politically defensible**. An attribution-centered micro-royalty model has the rare advantage of appealing to values across the ideological spectrum:

- **Fairness and Justice:** The proposal speaks to fairness by protecting creators from exploitation and ensuring those who add value are not left empty-handed. It can be framed as correcting a power imbalance: big tech firms reaping profits from artists' work should share some with the artists – a narrative of social justice and standing up for the “little guy.” It also aligns with the labor movement's ethos: just as workers deserve fair wages, creators deserve fair compensation when their labor (intellectual labor) is used. Many on the left champion the rights of gig workers and freelancers; one can view uncompensated training-data usage as analogous to uncompensated labor, which this model fixes in part.
- **Meritocracy and Property Rights:** From a more conservative or classical liberal angle, the model reinforces property rights and merit-based reward. It treats creative expression as a form of property or capital that, when employed by another (AI companies), should earn a **royalty** just as using someone's patented invention might. This will appeal to those who see creators as entrepreneurs or rights-holders entitled to the fruits of what they own. Micro-royalties also preserve incentive: they ensure that talented creators have reason to keep creating because even if AI uses their work, it drives income back to them. In a right-leaning view, this avoids the disincentive problem where people might stop creating if their work is simply taken with no reward. Instead, it's a market-driven solution: creators effectively license their work to AI via an automated royalty system, rather than relying on government subsidies.
- **Innovation and Growth (Business-minded Centrists):** Politically, one can argue that this model strikes a balance that allows AI innovation to continue (no outright bans or crippling fees), while also creating a **new market** for creative content in the AI age. It's akin to how the music industry adapted to digital streaming – by accepting lower per-unit revenue in exchange for new distribution channels. Here, creators accept that AI will use their work widely, but they gain a new revenue stream from every use. This could be sold as a win-win engine of growth: more content available

to AI (since creators feel safe contributing it) leading to better AI products, and a **creative economy** where even micro-payments, at scale, funnel money to potentially thousands of creators. Notably, experiments show consumers “*respond positively to fair compensation models*” for AI content (goizueta.emory.edu), which implies a competitive advantage for companies that adopt them (consumers might favor AI platforms known to be ethical and supportive of artists). Thus, even free-market advocates can see this as an efficient outcome where all parties trade value transparently.

- **Cross-ideological Appeal:** At a deeper level, attribution with compensation taps into shared societal values: **no one likes plagiarism** or unfair taking. Whether one’s hero is Adam Smith or Karl Marx, taking someone’s work without credit or payment feels wrong. This proposal institutionalizes a remedy in line with both **moral dessert** (people get what they deserve) and **economic justice**. It’s also flexible enough to adapt – if one faction argues the rates are too high or too low, those can be democratically adjusted without scrapping the whole system. In contrast, extreme solutions (like banning AI from using any copyrighted material) polarize by pitting innovation against creators; micro-royalties seek a **middle path** that most can agree on, much like the idea of a “creative dividend” on AI.

By highlighting these various principles, policymakers can build a broad coalition. For example, legislation to mandate or encourage attribution and micro-royalties could gain support from progressive lawmakers concerned about tech’s impact on workers, and simultaneously from more conservative lawmakers interested in protecting intellectual property and promoting enterprise (many artists and publishers are small businesses). The tech industry might support it as an alternative to heavier regulation, and the general public as a common-sense fairness measure. This broad appeal is critical, because the changes needed (e.g. new copyright exceptions or collective licensing arrangements) will require political will. **Framing matters:** one could call it an “AI Creativity Fairness Act” (emphasizing fairness) or a “Property Rights and AI Transparency Act” (emphasizing rights and transparency). Both narratives lead to the same mechanism. The key is that attribution-based compensation can wear many hats politically, making it a **resilient policy proposal**.

3.5. Implementation Roadmap and Evolution

When and how should attribution-based compensation be implemented? The answer: as soon as feasible, but in phases. **Phase 1 (Immediate–Near Term)** could involve voluntary industry adoption and pilot programs. Forward-looking AI companies might start adding attribution features now – indeed, some like OpenAI have hinted at tools that are able to provide citations in AI outputs. Likewise, forward-thinking content platforms (art sites, stock photo libraries, etc.) could partner with AI firms to license content for training with agreed attribution and royalty terms. These pilots, perhaps initially limited to certain domains (e.g. image generation with a known set of artists, or news article summarizers that credit journalists), would demonstrate the concept and work out kinks in technology and payment flows. During this phase, participation could be opt-in and the micro-royalty rates could be

nominal – the goal is to prove that the system works without breaking either the AI utility or the creator’s willingness to join.

Phase 2 (Mid Term) would aim for standardization and broader coverage. If Phase 1 shows success, there will be momentum to formalize attribution requirements, possibly through **industry standards or regulations**. For example, regulators might update copyright laws to explicitly require attribution for training-data usage (some lawsuits are already pressing in this direction (academic.oup.com)). Governments could mandate that generative AI systems above a certain capability include an “attribution mode” or logging mechanism. At the same time, collective management organizations might form to handle the sheer volume of micro-transactions – analogous to ASCAP/BMI in music or collective licensing for photocopying in publishing. In Phase 2, the priority class (professional creators) would be fully integrated, and the net could widen to semi-professional and independent creators. The remuneration might remain small per use, but with more outputs being attributed (for instance, AI chat assistants citing sources for factual answers, image generators listing artists), the flow of rewards would increase. We could see the emergence of an “**AI royalty statement**” for creators, much like streaming statements for musicians.

Phase 3 (Long Term) envisions a mature, possibly **universally applied system**. In this phase – say 5-10 years out – attribution might be so ingrained that even contributions from the long tail of creators are identified and rewarded. Advances in AI interpretability could make it routine to link any generated content back to influences, even those not originally registered. Laws may cement the practice globally (much like the Berne Convention did for moral rights). Compensation models could also evolve: micro-royalties might be supplemented by **collective funds** (e.g. a portion of AI company profits distributed to creator communities in proportion to attribution data) or by **tiered royalties** (higher rates for certain uses, such as commercial exploitation of AI content). The system should remain flexible: if at some point AI becomes so powerful that individual attribution is less meaningful (e.g. truly novel creations), policies can adjust to emphasize other forms of benefit sharing. But as long as AI relies on human-created data (which it likely will for the foreseeable future), attribution-based micro-royalties will continue to make sense and can only grow in coverage. An ultimate vision is an **AI training data registry**: a global database where all creators can list their works, with AI systems required (or strongly incentivized) to check against it and report usage. That registry could then automate royalty distribution at scale. It sounds ambitious, but it follows a precedent – consider how the collective management of music royalties worldwide, though complex, has been largely achieved through international cooperation.

One important consideration is **when to legally require compensation** versus relying on market adoption. A prudent approach is to allow industry some time to adapt voluntarily (to avoid stifling innovation with sudden mandates) but set a clear **deadline** after which certain practices become mandatory. For example, regulators might announce that in two years, any generative AI product on the market must implement basic attribution and contribute to a creators’ fund. This gives time to develop standards and technology while creating urgency. The evolution of the model should also be guided by data: as attribution logs come in, policymakers can assess who is benefiting, who is not, and adjust the mechanisms to ensure fairness (this could address, say, if big publishers are getting most payments but

freelance creators are still left out, one might tweak the formula or add supportive measures).

In summary, the implementation should **start now, but start small**, prove itself, and then ramp up via policy and technology improvements. By beginning with micro-royalties, we treat the current period (2020s) as a testing and norm-setting era, so that by the time generative AI is truly ubiquitous, the norms of attribution and compensation are already baked in. Delaying would only allow entrenched patterns of uncredited appropriation to solidify, which would be harder to reverse. The window of opportunity is open now. It is precisely the time to **experiment, legislate, and iterate** on attribution-based solutions. Over time, what begins as micro-royalties to a subset of creators could evolve into a **comprehensive framework of creative equity**, where generative AI is not a zero-sum displacer of human artists, but a collaborator that shares its gains with the creative community.

Attribution vs. UBI and Robot Taxes: Complementary, Not Redundant

The challenges of AI's impact on society have led to macro-level proposals like Universal Basic Income (UBI) or "robot taxes" (taxing AI/automation to fund social welfare). While such economic measures may be part of the solution space, this declaration contends that **attribution fulfills a distinct and irreplaceable role** in the AI era. It addresses dimensions of human need and social value that pure monetary redistribution cannot.

Attribution Addresses Meaning and Recognition: UBI, in concept, provides people with money to meet their basic needs in a post-work future. But as psychological and philosophical research makes clear, *human beings seek more than sustenance – we seek meaning, purpose, and recognition*. Simply receiving a government check, while helpful for survival, does not tell a person that their creative talents are valued by society. Sociologists warn that separating income from contribution can undercut the social recognition that traditionally comes from work or creative output ([linkedin.com](https://www.linkedin.com)). In contrast, attribution explicitly ties recognition to contribution. It publicly affirms that *"this person's creativity mattered here."* This kind of recognition has profound effects on self-worth; it satisfies the esteem needs that sit high on Maslow's pyramid. As one basic income advocate noted, even with UBI, people will still "equate gainful employment (or productive contribution) with dignity and recognition" (basicincome.org). So if AI displaces creative jobs, UBI might keep artists fed, but without attribution they lose the dignity that came from having their name on their work. Attribution ensures that creators retain **creative legitimacy** – society still sees them as the authors of something valuable, even if AI helped disseminate or transform it.

Attribution Fosters Ongoing Creative Agency: A robot tax might funnel money into public coffers that could support unemployed artists or fund public art. That is commendable, but it treats creators more or less as beneficiaries of **charity or social welfare**, rather than as active participants in the economy. Attribution-based compensation, on the other hand, keeps creators directly **engaged**. When an artist receives micro-royalties because her style was used by an AI, she is being rewarded *as a creator*, not as a passive citizen. Her relationship to the AI is one of a licensor or collaborator, not merely a victim made whole by the state. This is crucial for **agency**. Philosophers of work (like Marx) argued that being

alienated from the product of one's labor is detrimental to one's humanity. UBI without attribution might lead to a scenario of alienation: the artist's work fuels AI outputs, but she is alienated from that process and just gets a stipend unrelated to her creative impact. Attribution reconnects her to the product: her name and a portion of value travel with her contribution. It thereby mitigates alienation by **integrating human creators into the AI value chain**. Even if the money were equivalent, many creators would prefer to earn through recognition of their work rather than through an unrelated stipend, because the former validates their *creative identity*.

Moral Rights and the Intangible Value of Credit: As discussed earlier, international law and ethical theory recognize *moral rights* that are separate from economic rights ([humanrights.com](https://www.humanrights.com)). A robot tax or UBI scheme deals purely in economic outcomes – it might compensate for job loss or general welfare, but it does nothing to preserve the *moral right of attribution*. Failing to credit an author is considered a harm in and of itself, injurious to the author's honor or reputation (en.wikipedia.org). No amount of money can fully erase that harm, because it's not about money – it's about **respect**. In fact, one could imagine a dystopic scenario where AI companies pay a hefty tax to government and thus feel absolved while they erase authors' names from all content; society might be materially okay (tax money funding public goods) but culturally impoverished and unjust in the treatment of authors. The declaration's vision insists that we must not trade moral rights for money – we need both material and moral amelioration. Attribution is the way to secure those moral rights. It provides something that UBI cannot print: **authorship credit** and a continued place for humans in the lineage of ideas and creations.

Social Cohesion and Public Buy-In: There is also a pragmatic angle – public acceptance of AI might depend on more than just economic adjustment. If people see AI as annihilating human creativity and giving nothing in return except maybe a government check, resentment and cultural resistance will grow. Humans have a need for **story and identity**; we cherish the idea of authors, artists, inventors – it gives a human face to achievements. Attribution keeps that human narrative alive in each AI-generated story or artwork by saying “inspired by X” or “based on Y's work.” This not only benefits creators but also the public, which can still connect creations to human stories and communities. UBI cannot offer that narrative; it operates in the background. In short, attribution humanizes AI outputs, which is vital for maintaining **social trust in technology**. The populace is more likely to embrace AI if it is seen to elevate human creators rather than replace or erase them. Robot taxes might redistribute wealth but do not by themselves confer *meaning*. Recognition and attribution are about distributing **meaning and honor** in society, not just wealth.

Complementarity, Not Competition: It's important to emphasize that this is not an either/or choice. Attribution-based micro-royalties and UBI/robot taxes can and perhaps should coexist. Attribution is a **targeted solution** for those who contribute to AI's capabilities (creators, data providers), while UBI is a **universal solution** for society at large in case AI upheaval affects jobs broadly. They address different layers of the issue. Indeed, if AI-induced productivity is so massive that a robot tax or UBI is viable, integrating attribution will only strengthen the ethical case for such redistribution by ensuring we haven't trampled individual rights on the way. Conversely, if UBI ensures everyone's basic economic security, attribution still plays a critical role in *allocating prestige and maintaining motivation* for creative endeavor. A painter might have UBI to live on, but it's the attribution in AI that will

inspire her to keep painting, knowing that even if AI uses her style, her name and a slice of earnings travel with it.

To summarize, **UBI/robot tax and attribution address fundamentally different human needs**: the former addresses subsistence and macro-distribution of AI's gains, the latter addresses identity, justice, and micro-distribution to actual contributors. Economic redistribution is not a substitute for moral recognition. A humane future with AI will likely require both: *material support for all* (so no one is left destitute by automation) and *recognitional support for contributors* (so creators continue to create and feel valued). Attribution is how we ensure the latter. It is the piece of the puzzle that keeps our solutions human-centered, not just economically centered. By implementing attribution and micro-royalties, we acknowledge that people are not only economic units but also **moral agents and creatives** whose dignity must be preserved even as technology advances. In doing so, we uphold a richer conception of justice – one that speaks to the soul as well as the stomach.

Conclusion: Toward a Human-Centric AI Future

This declaration has argued for an **AI Attribution** paradigm as an essential pillar of ethical AI deployment. Backed by insights from neuroscience, psychology, philosophy, and law, we have shown that attribution is far more than a courtesy; it is a recognition of human dignity, creativity, and **moral rights**. It offers a concrete path to address the social challenges of generative AI by re-balancing the relationship between human creators and AI systems. Attribution, especially when coupled with a micro-royalty compensation scheme, provides a framework where AI innovation and human creativity can enter a **symbiotic relationship** instead of a zero-sum game ([createdbyhumans.ai](#)). In this envisioned future, every time an AI model generates content that echoes human artistry or knowledge, it does so in a way that **honors the human source** – displaying their name, and even delivering a token of economic value to them. This transforms AI from an appropriator to a **collaborator**. It ensures that as AI prospers, so do the people who feed it with culture and information.

Of course, attribution and micro-royalties alone will not resolve every concern about AI. They will not immediately answer hard questions about AI accountability for errors, deepfake misuse, or the need to retrain workers. Those issues require additional policies and innovations. But without attribution, any solution will be built on a shaky ethical foundation. We risk a future where human creators are disenfranchised in spirit even if provided for in body. By implementing attribution, we anchor the AI revolution in respect for the individual and the intrinsic value of human creativity. We also unlock practical benefits: a transparent provenance for AI outputs (aiding trust and quality), and a sustainable incentive structure for creators to keep contributing new works (so that AI's well of inspiration never runs dry) ([createdbyhumans.ai](#)). Indeed, if we want AI to continue learning and improving, we must ensure *humans remain incentivized to produce great art, writing, research, etc.* – which they will only do if they are recognized and can make a living. Properly aligning these incentives through attribution is thus not only a moral choice but a pragmatic one. In implementing this declaration's vision, we call on multiple stakeholders: **AI developers** should build attribution tracking and output citation features, and embrace licensing models that include micro-royalties; **policymakers** should update laws to require source transparency and to

facilitate collective micro-licensing of works for AI, bridging the gap between copyright and AI technology (academic.oup.com); **creators and publishers** should organize to negotiate fair terms, possibly creating registries of works available for ethical AI use; and **users** should demand and favor AI systems that respect attribution – using their power to shape market preferences. There is an optimistic alignment here: what benefits creators (attribution) can also benefit users (through greater transparency and richer cultural context) and even AI companies (through access to higher-quality, properly licensed data and avoiding litigation) (goizueta.emory.edu; publishers.org). The declaration’s proposal thus charts a path forward that is **constructive and cooperative**, rather than adversarial.

In closing, let us remember that **human dignity and creativity are not relics to be discarded in the age of AI, but treasures to be carried forward**. The goal of technology should be to amplify human potential, not eclipse it. By enshrining attribution as a core principle, we ensure that the **human face of creativity** remains visible in every algorithmic product.

Through attribution and fair compensation, generative AI can become a true symbiosis – drawing on the depth of human experience while respecting and rewarding the people behind that experience.

This declaration presents a vision of an AI-powered future that does not forsake the individual creator. It is a future in which an **AI Attribution Framework** helps society navigate the upheaval of generative AI with its values intact. It is, ultimately, a call to action: to academics, to lawmakers, to engineers, to artists – to come together in crafting policies and systems that **recognize human genius in even the AI’s most dazzling feats**. By doing so, we safeguard human dignity, sustain creative livelihoods, and foster a culture where humans and AI thrive together in mutual respect. The challenges are great, but so is the prize: a future where technology enriches humanity without erasing its authors. Let us begin, with attribution as our compass, to build that future.

Anticipating the Critics

Attribution will raise objections—technical, philosophical, and political. We welcome critique and respond with clarity:

Objection	Our Response	References
-----------	--------------	------------

“Probabilistic models can’t trace influence.”	True—but partial attribution is possible using tools like influence functions, data ablation, and Shapley value approximations. In law, epidemiology, and social science, probabilistic causality is sufficient to assign responsibility and reward. So too in AI.	<ul style="list-style-type: none"> - Koh, P.W., & Liang, P. (2017). <i>Understanding Black-box Predictions via Influence Functions</i>. arXiv:1703.04730 - Lundberg, S.M., & Lee, S.I. (2017). <i>A Unified Approach to Interpreting Model Predictions</i>. arXiv:1705.07874 - Oberdiek, H. (1977). <i>Causality and Responsibility</i>. SpringerLink
“All public data is fair use.”	Legal access ≠ ethical legitimacy. Cultural visibility is not consent. Attribution transforms exposure into acknowledgment and turns exploitation into structured reciprocity.	<ul style="list-style-type: none"> - Lin, P.K. (2021). <i>Fair’s Fair: How Public Benefit Considerations in the Fair Use Doctrine Can Patch Bias in Artificial Intelligence Systems</i>. SSRN:4116483 - Newton-Rex, E. (2023). <i>A New Nonprofit Is Seeking to Solve the AI Copyright Problem</i>. Time
“This restricts innovation.”	Innovation without consent or credit is exploitation at scale. Attribution enhances social license, data quality, and public trust. Long-term innovation depends on legitimacy, not just speed.	<ul style="list-style-type: none"> - Berger, V. (2025). <i>The AI Copyright Battle: Why OpenAI And Google Are Pushing For Fair Use</i>. Forbes - Vuyyuru, H.K. (2024). <i>Charting the Future of Responsible AI Innovation</i>. Analytics Insight
“This reinforces neoliberal property logics.”	Attribution is not ownership. We support pluralist models: symbolic credit, commons-based recognition, data cooperatives, and relational accountability. This isn’t commodification—it’s ethical memory.	<ul style="list-style-type: none"> - Donway, R. (2023). <i>Neoliberalism on Trial: Artificial Intelligence and Existential Risk</i>. Econlib - Morozov, E. (2023). <i>Can AI Break Out of Panglossian Neoliberalism?</i>. Boston Review
“It’s impractical at scale.”	So were copyright, Creative Commons, citation indexing, and GitHub. Attribution will scale through protocols, automation, and open standards—and will become machine-readable by default. Not easy—but feasible.	<ul style="list-style-type: none"> - Longpre, S., et al. (2023). <i>The Data Provenance Initiative: A Large Scale Audit of Dataset Licensing & Attribution in AI</i>. arXiv:2310.16787 - MIT Media Lab. (2023). <i>Data Provenance for AI</i>. MIT Media Lab

“It won’t scale to trillions of tokens or images.”

Full traceability for every token is infeasible—but layered, probabilistic, or statistical attribution is tractable. We don’t need 1:1 tracing. We need signal-based accountability, much like royalty models or citation networks.

- MINT-1T: *Scaling Open-Source Multimodal Data by 10x*. [Unite.AI](#)
- Young, M. (2023). *MINT-1T: Open-Source Multimodal Dataset Scaled to One Trillion Tokens, Enabling More Capable AI Models*. [Dev.to](#)

“Attribution will be gamed or spoofed.”

Like any incentive system, it’s gameable. But cryptographic provenance, model audits, and reputation layers can reduce fraud—just as in finance, authorship, and intellectual property.

- Harvard Law School. (2023). *Did ChatGPT really say that?: Provenance in the age of Generative AI*. [Harvard Law](#)
- Collomosse, J. (2024). *To Authenticity, and Beyond! Building Safe and Fair Generative AI upon Digital Provenance*. [IEEE Computer Graphics and Applications](#)

“It will slow down model deployment.”

Perhaps—but like carbon audits, clinical trials, or ethics review boards, that’s a feature, not a bug. Friction invites reflection. Good governance is never entirely frictionless.

- Forbes Technology Council. (2024). *The AI Model Slowdown: Don’t Be Alarmed*. [Forbes](#)

“Users don’t care where data comes from.”

Maybe not all—but creators, artists, scientists, and communities do. They are organizing. Attribution is not just about individual rights, but about cultural memory and resisting epistemic erasure.

- Anadol, R. (2025). *I make millions from AI art—but the law has to be fair* ([The Times](#))