VERMILLIO

U.S Copyright Office

Notice of Inquiry on Artificial Intelligence & Copyright (Dkt. 2023-6)
Comments of Vermillio, LLC
Dec 6, 2023

Who is Vermillio?

Vermillio is the creator of the first generative AI platform built specifically to protect the work of content creators. Vermillio's platform enables IP holders to leverage authenticated generative AI: where the authenticity and ownership of all synthetic content generated via Vermillio's platform is tracked seamlessly and transparently. We call this "Authenticated AI". Vermillio was founded on the belief that generative AI presents a huge long- term opportunity for IP owners and content creators, so long as those systems and frameworks are fully consent-driven and authenticated. Vermillio builds these Authenticated AI systems and frameworks, and makes them available to the world.

Everyone agrees that content creators' rights ought to be protected - the question presented by generative AI is what, exactly, the scope of that protection ought to be. Everyone agrees that technological innovation is desirable - the question is, at what cost. While others have presented arguments to the Copyright Office suggesting that striking the proper balance between innovation and creators' rights is simply too difficult to manage as a practical matter, Vermillio respectfully disagrees. Our platform proves that such a balancing of interests is not only possible, but is already readily available in the marketplace. Users of Vermillio's platform - including studios, music labels, gaming companies, actors, singers, and brands - are able to leverage the full spectrum of benefits afforded by generative AI, while also ensuring that the fundamental rights of content creators to consent, compensation, and credit are protected. Authenticated AI, as exemplified by platforms like Vermillio's, empowers IP owners and creators of all stripes to offer fans new, immersive ways to engage with their favorite content through generative AI while ensuring that content creators' rights are respected.

Vermillio was founded by a team of technologists with over a hundred years of collective experience building AI software and scaled transaction systems. Since the company's founding, it has worked in partnership with best-in-class IP holders and content creators to develop its proprietary Authenticated AI technology. Vermillio launched its first consumer-facing engine earlier this year. In partnership with a well-known movie studio, we enabled over [1.5M fans to create one of a kind portraits of themselves in the style of one of the world's most beloved franchises](#). Alongside an international record label (and with enthusiastic support from one of the label's legendary musicians and songwriters), Vermillio [allowed fans to "remix" the official album art and create new, one of a kind tracks based on excerpts of the music from the album.](#) Each of these creations was generated with an engine that was trained on actual,

properly-licensed content from the IP holder, and all outputs generated through the system were authenticated and tracked using Vermillio's proprietary platform and Trace ID.

What is Vermillio and Trace ID?

Vermilio enables IP holders to track the way their IP is used, so that permitted creators and app developers can use the IP to build consumer experiences that the IP holder has specifically consented to. The Trace ID™ forms the basis of Vermillio's Authenticated AI experience. Vermillio leverages the benefits of blockchain technology (along with digital signatures) to ensure that IP Holders can track and control in real time who is permitted to access their IP for training, or to develop synthetic content. This approach has multiple benefits: not only does it ensure that access to the IP holder's content for the purpose of training generative AI models is appropriately restricted, it also ensures that consumers or content creators who leverage that IP to develop synthetic content are doing so in a fully licensed and approved environment. Vermillio's use of "smart contracts" - software that runs on blockchain, and that enables near-instantaneous tracking and management of on-chain data - allows the platform to evaluate any piece of synthetic content generated via the platform and determine whether that content was based on IP that was properly licensed for the purpose. Each piece of original content (ingested by the platform), as well as each piece of derivative content created when using Vermillio, is assigned a unique Trace ID. Essentially, Vermillio is a quick way to resolve IP disputes, answer rights lineage questions, and verify authenticity. The platform does all this while maintaining the privacy of the IP holder and the consumer/content creator.

Vermillio protects IP holders' content through every stage of the AI data lifecycle (including input, training, and output). We begin tracking the content as soon as it enters Vermillio's ecosystem, and continue to track the data and metadata throughout the atomization, training, and model creation stages using unique Trace IDs. All synthetic content derived from the licensed IP is similarly tracked, and attributed to the IP from which it originated.

We developed the platform to address the following use cases:

- IP holders who want to manage and monetize their IP;
- App developers who want access to properly-licensed training data, models, and/or content from IP holders; and
- Platforms (such as Facebook, TikTok, YouTube, Instagram, X, Spotify, DSPs) that want to enable content authentication.

IP Holders. As a result of taking an "opt-in" approach to training generative AI models (as opposed to simply using all available IP on the Internet without permission to train the model), Vermillio knows all the parties involved in leveraging the IP on its platform. The platform has also been built to enable third parties to track and manage payments for IP usage, allowing IP holders (and their associated content creators) to be compensated properly for their work and for any synthetic content that derives from it.

App Developers. App developers can use Vermillio to license content into their generative AI-based applications. The content can be used in any way the IP holder and app developer agree upon (e.g., in a game, or in an app that creates new synthetic content). The app developer calls to the platform when the IP holder's content is used, and when new synthetic content is created, the developer's application automatically integrates that synthetic content into the platform to be tracked and traced in the same way the IP holder's original content would be.

Platforms. Platforms (social media, DSPs) that integrate Vermillio through APIs can provide a distribution point for Authenticated AI while being assured that the content is both approved by the owner and credited properly for payment purposes. If content is suspected of being unauthorized synthetic content, then Vermillio allows the platform to automatically check the content against the Trace ID and enable the content holder to determine the best course of action. This relieves the platforms from the obligation to self-police, and allows the benefits of Authenticated AI to be shared with millions of users.

Verification of Authenticity vs. Policing

Other commenters have suggested that the only way to manage the explosion in infringing synthetic content is to rely on existing copyright enforcement mechanisms. While these enforcement mechanisms are useful in principle, in practice for IP holders the effort amounts to a never-ending game of Whac-A-Mole. Fighting infringement after the fact has quickly become an endless, Sisyphean task, pitting an IP holder's limited resources against the exponential growth of infringing synthetic content on the Internet. We can and should do better. We built Vermillio - with its ability to verify and authenticate inputs and outputs in real-time - in an effort to solve this thorny problem. Using a combination of on-chain data and digital signatures, Vermillio can validate content that has been fingerprinted, and can use detection AI to determine if a particular piece of IP is being managed by the platform. Vermillio also provides mechanisms and APIs for an IP holder to revoke access to an asset that was not properly licensed.

To be clear, while we expect that Vermillio's approach to Authenticated AI will dramatically reduce infringements, the Trace platform is not an active policing system. As noted above, policing content is a costly race against ever-improving AIs and GANs. Our focus is on helping ensure appropriate monetization for IP holders whose content is being leveraged by generative AI, expanding the proper, licensed use of IP for generative AI purposes, and tracking lineage of IP rights (both of the original input, and all synthetic output) publicly on-chain.

Soft Binding vs. Hard Binding

Distribution platforms have introduced techniques to help identify content on their platforms that might be infringing (e.g., Content ID). These software programs are trained on the original content from the IP holders, and are designed to identify exact replicas of the content. . As a result, these programs are called "hard binding", as they are looking for one-to-one matches and are deterministic in their outcomes. For obvious reasons, "hard binding" techniques do not tend to work well in a generative AI context. Hard bindings lack

the ability to find similar, but not identical, content. New techniques that leverage robust content similarity are required. Generative AI creates derivatives of the original content (training data) rather than one to one matches so any binding technique needs to focus on an ensemble approach that leverages syntactic (similarity without meaning) and semantic (similarity with meaning) bindings, otherwise known as "soft bindings" (Trace ID's). Soft bindings are powered by digital hashes and fingerprints. Derivatives of the original contain traces of those fingerprints allowing them to be tied together. Soft binding approaches will drive the consent and compensation rigor for this opportunity, and allow for it to scale with benefit to all.

Why Authenticated AI?

We started to develop Vermillio over 3 years ago because we saw a shift that reminded us of the shift from analog to digital distribution. When studios, music labels, gaming companies, actors, singers, brands and content business leaders chose to move from analog formats to digital formats like streaming, they gained massive benefits and efficiency of distribution.

In the music industry (as one example), that shift played out with the launch of Napster, and eventually Spotify. Since then, streaming services have demonstrated that most consumers would rather pay $10 to $20 a month for the convenience of unlimited, personalized music over free (but pirated) alternatives. The music industry's shift to a subscription model had other, unexpected consequences, including increasing exposure to and benefits for artists. Many little-known artists who would otherwise never have secured a record deal benefited from the streaming revolution, finding new audiences for their work — some have even gone on to stardom, thanks to the distribution power of platforms like YouTube.

All these benefits were predicated upon a critical (if, at the time, less than fully conscious) choice impacting the futures of content creators and owners alike: to allow social platforms to digitize the creators' content in pursuit of users. Turning hundreds of years of human creativity into 1s and 0s allowed these new algorithms to parse the content and provide it to audiences in the most precise and efficient manner possible. While this new model gave the social platforms incredible leverage (because machines were highly effective at distributing the content), the power of creation remained in the hands of humans.

We believe that we are at a similar inflection point with the rise of generative AI. Today, however, not only have the means of distribution been digitized, but if left unchecked the power of creation will soon rest primarily in the hands of machines. The creative act - everything from filmmaking, songwriting, creative writing, to visual arts, all critical aspects of the building blocks of culture – can and will be synthesized, digitized, and commodified.

We use the term "synthetic content" as a catch-all term for media created by machines using generative AI, along with training data that we call a "digital signature". The digital signature is simply the 1s and 0s that represent the content digitally, and what a computer reads at an atomic level. The digital signature can come from existing IP such as a movie, artwork, voice, or song, or from a wholly new piece of creative art made by AI based on patterns discerned from

existing datasets. For example, the 1s and 0s that describe part of a Jay-Z song in digital form can be used to create synthetic content where Jay-Z appears to be performing any song (even songs that are not his own). This new creation is not a mash-up of actual Jay-Z recordings— it is an entirely new creation, based on a small sample of his digital footprint.

With the rise of synthetic content and the murky legal waters surrounding it, content creators and IP holders have to decide:

i) Whether to embrace an emerging status quo where unauthorized third parties will train generative AI models on the IP holders' proprietary content without consent or compensation, and use synthetic content output that these models create however they wish, paying nothing for it; or

ii) Whether to chart a new path via Authenticated AI where existing technologies, including AI and blockchain, are used to allow third parties to buy access to an authorized digital signature of training data that is digitally controlled by an IP owner.

We see a massive opportunity in the second path, and would propose the following set of fundamental values to ensure that this opportunity is made available to all, and not just a few technology companies.

Fundamental Values for a Healthy, Vibrant Marketplace Leveraging Generative AI

Over the last decade, the digital transformation was, among other things, a race for users. Technology companies dominated this transformation by controlling both the users and their data, while at the same time avoiding accountability for copyright infringement on their platforms.

This dynamic led to both massive wealth generation (concentrated in one industry), as well as marketplace domination (concentrated in that same industry). When content companies and advertisers insisted on some level of policing of IP infringements, technology companies developed their own tools (e.g,. Content ID), but each platform developed its own tools and guidelines with content owners having very little say. While technology companies policed themselves, content owners were left with little choice but to accept bad outcomes, or miss out on the incredible commercial opportunities that access to those companies' platforms provided.

This "safe harbor" protection against claims of copyright infringement was intended to support innovation (which it did), but it ultimately stifled the economy of IP holders. Artists, filmmakers, major talent and content owners were the ones who ultimately made these platforms successful and engaging, but revenue splits did not reflect the creators' contributions to user acquisition, engagement, and retention. Hundreds of billions in asset value flowed to the technology platforms, while content owners received pennies on the dollar. The advent of generative AI is an opportunity for us to balance those scales by better aligning incentives, protecting IP holders, providing tools for Authenticated AI experiences, and creating opportunities for all parties to profit.

To build this future, we first need to agree that the scales ought to be more finely balanced between content owners and technology platforms. Companies like Vermillio, with platforms like the Trace platform, can assist in this regard by providing Authenticated AI options for IP owners, distribution platforms and fans.

Third Party Authentication

Third party authentication tools like the Vermillio platform and Trace ID's allow for three key developments in the marketplace:

i) Allowing social platforms and DSPs to maintain their arms' length relationship to content creation, and continue to enjoy their DMCA "safe harbor" . As synthetic content increases exponentially, content platforms will need meaningful content moderation and authentication tools from a variety of companies that are dedicated to improving their effectiveness.

ii) Giving content owners and creators the ability to control their content, and creating a marketplace for content authentication and protection. This competition will drive innovation. We've seen the rise of similar technologies to protect brands inside the major platforms. When brands only wanted their advertisements to show up next to trusted content, they insisted on the implementation of third party systems for authenticating content so that their ads only showed up alongside approved content. Today, these types of systems authenticate billions of impressions every minute. Why can't this be the case for synthetic content, as well?

iii) Encouraging the growth of an "app store"-like marketplace for generative AI startups. Authenticated AI will lead to the development of robust, specialized engines that will benefit both content owners and technologists. We already see this developing, with announcements like Meta's virtual assistants with well-known celebrities.

With these three key developments, we see the potential for generative AI to bring economic upside to a more diverse set of business stakeholders. For this marketplace to grow and for all parties to profit, however, we believe the following values must form the foundation:

- Consent and Control
- Transparency
- Clear Credit, and Fair Compensation

Consent and Control

We believe that IP holders should control how and when their content is used to train AI. This would allow IP holders to take advantage of the Generative AI opportunity while maintaining control over their IP and license their content for training and development by approved partners. Creating new revenue streams while protecting the rights of IP holders, Authenticated AI would limit foul play with their IP in the form of deep-fakes or other unauthorized content. Undesirable outcomes could be prevented by using Authenticated AI technology to highlight synthetic content created without the seal of approval of the person or asset featured.

Transparency

We believe that fans want to interact and create with generative AI. We have already seen it when launching our first consumer-facing experiences. At the same time, audiences also want to know when they're interacting with an AI. They do not want to be fooled, or have a lower-quality experience because the AI was trained without the proper data or without the permission of the IP holder.

Clear Credit and Fair Compensation

We believe that artists, musicians, filmmakers, writers and creators of all kinds should receive credit when their work is used to train an AI model, or when their work is used as a basis of any synthetic content. This will lead to more fans engaging with their original IP, and a positive collaboration between all parties. IP holders should be compensated for the use of their content.

Conclusion

There are ways to more effectively balance the needs of content creators against the desire to support continued technological innovation in the AI space. Authenticated AI is an existing technology that can create opportunities for IP holders to monetize their content, allow third parties to train their generative AI models on that IP, and enable consumers to create and enjoy synthetic content, all while helping to reduce the likelihood of infringement.

Authenticated AI would encourage citizen creators and companies looking to access the long-tail of existing content to open up their back catalogs, in the same way that eBay created a marketplace for buyers and sellers of every object on earth.

Authenticated AI will enable studios to create marketing campaigns that have thousands of derivations, each personalized to the target consumer.

With Authenticated AI, a famous actor that defines a franchise will no longer have to come in and do voice edits. That same actor can be recreated over and over again, without ever having to step foot in a studio. Even the estates of deceased actors, singers and writers can now continue to monetize:a synthetic Sinatra singing the latest James Bond theme is now possible.

Authenticated IP can enable a vibrant makerplace where creators, technologists and consumers all benefit from generative AI, and where the scales of risk and reward are more balanced. Immense economic value will be created, all while protecting the artists, filmmakers, writers and creatives who define our culture, and reflect what makes us human.