This document is intended as a response to portions of Google LLC's comments in reply to the USCO's notice of inquiry, *Artificial Intelligence and Copyright*. More specifically, it is a response to the portions concerning the doctrine of fair use as it relates to training generative AI systems.

Google begins by positing an inherent competition between the "interests of authors … in the control and exploitation of their writings", and "society's … interest in the free flow of ideas, information and commerce". An ideal copyright system is able to balance the interests of both parties in order to for the public to best benefit from the labours of authors. This is undoubtedly true. But the appropriate balance is not self-evident – it may be argued a determination that the ingestion of creative works is not automatically fair use is in the genuine interest of the public. An overly permissive regime with respect to the ingestion of creative works would profoundly disincentivise the open sharing not only of works completely authored by a human, but also works partially crafted by machine. If anyone's labour can be instantly snatched up and iterated upon without compensation or credit in enormous volumes at inhuman speed, you would need to closely guard it, lest it be rendered worthless. I have even begun to witness people using AI guard their prompts, fearing someone else may come along and appropriate their effort for their own ends. The production speed and volume of rip offs produced from an original work has historically been limited by the human factor. An author sharing their work openly for the world to enjoy would invariably make the determination that having their work out in the world outweighed the risk of it being ripped off. It used to take genuine time, effort and often skill to create derivatives. In a world where this is no longer the case, where it really is as easy as pressing a button to receive a half-dozen variations on an original work, or works in the style of someone's oeuvre, authors of all stripes, users and non-users of AI alike, will increasingly retreat into walled gardens to protect what little value their works may accrue.

In this regard, careful guardrails around the ingestion of works for training generative AI are, contrary to what Google claims, an exemplar of copyright's intended purpose. Authors would be free to share their work as they please, safe in the knowledge that it cannot be exploited non-consensually in ways that would undermine the value it holds. Guardrails would not only promote the continued sharing of creative works by authors of all kinds, they would also encourage the growth of licensing markets that see these revolutionary AI tools developed quickly and ethically with consideration for all stakeholders. It would genuinely be a win-win scenario.

The boundary of fair use is determined both by ideals – a determination of what "fair" really means – and by pragmatic concerns – how best to promote creative expression and allow the society at large to reap the benefit of that expression. As historically crucial factors (like human limitations on the competing production of derivative or stylistically derivative work) become moot, the boundary of fair use must change in kind. To ignore these shifts in technology and circumstance would render the "fair" in fair use an anachronism.

Google also stresses that machine learning systems capture only the "statistical relationships among training data" in an attempt to cast anything statistical in nature as unprotectable under copyright – an "underlying idea, theory or fact expressed in a work" as distinct from the specific work itself. This is a category error. We can see this by appealing to

a common issue in machine learning systems – overfitting. Overfitting occurs when an ML system accurately models the training data, but is unable to generalise beyond that or adapt to new data. In the case of a large language model like GPT-4, this may mean it memorises passages from a book contained within its training set. Why is an ML system, supposedly only designed to capture the unprotectable content of a work, able to capture verbatim unambiguously protectable content? It is because, when an AI system outputs in the same modality in its training data, the statistical inferences it makes also capture the protectable expressions within a work. When you are capturing statistical relationships, the protectable and unprotectable are not cleanly separable. Indeed, statistical relationships can capture all the protectable expression within a work just as well as the unprotectable expression. The relative positions of colours and forms in an artwork as modelled with statistics, may well represent the protectable expression within that artwork.

Google claims that fair use as a doctrine is not concerned with the specific type of market harm that generative AI would produce, citing the Second Circuits decision in *Authors Guild, Inc. v. HathiTrust.* "The fourth fair use factor 'is concerned with only one type of economic injury... the harm that results because the secondary use serves as a substitute for the original work.'" Key to Google's position is a specific understanding of what a "substitute for the original work" entails. The Second Circuit's example in their judgement gives the paradigmatic example of a book review. Directly from their brief:

"Book reviews often contain quotations of copyrighted material to illustrate the reviewer's points and substantiate his criticisms; this is a paradigmatic fair use. And a negative book review can cause a degree of economic injury to the author by dissuading readers from purchasing copies of her book, even when the review does not serve as a substitute for the original. But, obviously, in that case, the author has no cause for complaint under Factor Four..."

Crucially, they identify that a book review is not a substitute for the original work. It cannot be reasonably said that someone desiring to read a book would consume a book review in lieu of an actual book. But what of the output of a generative AI system build to produce books? A person desiring to read a book may well read an AI-generated book over another work that the AI system is mechanically dependent upon. In this sense, the outputs of a generative AI system, distinct from book reviews or discriminative AI systems that merely label or provide word analyses of books, *can* reasonably be considered substitutes for the original book. Being substitutions, the Fourth Factor comes into play. There is an inescapable fact here that generative AI systems are built using copyrighted works to produce outputs that *do* compete directly as substitutes in the same market for the original work. They are not merely works "of the same type" (e.g. books and book reviews are both formed from text) – they are in direct competition. The outputs of a system built from original books for the express purpose of producing new books (particularly when prompted in specific styles) will compete in a market as substitutes for the original works upon which the system is built.

But an overly-permissive regime with regard to generative AI would also cause market harm in the sense that it would undermine any fledgling market for the licensing of works for training of generative AI. If AI tools really are the future, then robust licensing markets will

be essential in ensuring that creatives are able to continue authoring creative works. They will also be essential in ensuring the continued existence of a human cultural “ground truth” for generative AI to reference, something of particular concern to AI developers themselves. Without a continued supply of human-generated data, future AI systems may see their quality and relevance decline.