Anand Desai's comments regarding artificial intelligence and copyright in response to the U.S. Copyright Office's August 30, 2023 Notice of Inquiry (88 FR 59942)

I'm a U.S. citizen writing only for myself with no significant financial interest in these issues. I have a very general understanding of generative AI such as ChatGPT and DALL-E (and no deeper knowledge of any other kind of AI).

Going by the question numbers of the Notice:

1. Risks and benefits of generative AI. AI could be extremely useful and a long-term qualitative advantage in practical and judgment oriented pursuits, the overall value of our economy, and – somewhat relatedly – our national security. As I understand it's a big next step from internet search engines in combining apparently relevant source information into a broad and logical narrative for the user.

AI also seems likely to be at least monetarily valuable for appealing information presentation and artistic type work that conjure basic emotions without much thought in the process. For instance, cheerful product descriptions, background music, and maybe eventually simpler TV shows. It could also could be a tool toward expressing more significant ideas.

Drawbacks that come to mind include career disruption from "de-skilling" the work of presenting and teaching well-known ideas, possible loss of market value particularly to lower quality works which AI could replace adequately and maintain wonderfully and to the businesses that create them (but some of these people could work with the AI to make a better product), possible loss of inspirations that come about in a more from-scratch process (which likewise must be weighed against the tool's tremendous help with background research solving tougher problems and marketing the result), an overall decline in real or perceived "authenticity" of art and published opinion (although this could be more an issue for "moral rights" and trademark-dilution concepts); and the potential to mislead the public by making convincing-looking mistakes, lies, or misattributions, biasing input data (perhaps not out of malice), or through a small number of dominant AI systems amplifying particular perspectives through inadvertence.

2. Industry perspectives.

Legal industry

AI would help the practice of law more efficiently serve the public. The practice of law involves a great deal of researching and analogizing from a history of arguably important points of reference, such as court opinions in cases involving arguably similar concepts. Virtual "teaching assistants" would make required educational programs more complete (especially in smaller markets) and less staff-intensive, decreasing often-taxpayer-subsidized tuition costs, and likely improving admissions "equity" concerns by making all schools quite capable and affordable whether one ultimately becomes a lawyer or "just" another kind of knowledgeable professional (currently most states require not just passing a bar exam, but graduating from a school whose accreditation is conditioned on most of the other students passing that exam – reducing the bill students could be stuck with would eliminate that as an excuse for prejudging eager but less obviously likely to pass applicants). Increased supply at lower cost would improve access to legal assistance for the public, and AI would help clients get reasonably well informed research for few hours of work. AI could also reduce "industry capture" of regulators by ensuring an ample supply of capable lawyers in all specialties for the government and aggrieved

individuals as well as wealthy insiders: currently only a few schools cover complex business tax issues in depth, for instance.

But, if rules such as copyright allowed organizations to exclusively efficiently exploit the concepts from large bodies of source materials – for instance, if large law firms or consultancies could cross-license among themselves access to tens of thousands of briefs they and their predecessors filed in court, and/or if jurisdictions other than the federal government prevented fair access to their rules – AI-driven increasing returns to scale could contribute to monopolistic abuse of employees and clients.

Good judgment and quality control remain paramount – and now to protect against errors created by one's own research tools too.

Same goes for other learned professions, I suppose.

Tax
Copyright coverage for use in AI could make cross-licenses more important in business. One concern I have is that they could create oligopolies. Another issue is that their tax treatment may not be clear. See IRS Notice 2006-34.

Free Software and Writing Projects
"Copyleft" licenses like the GPL and Creative Commons use copyright rules to ensure a body of work remains available to the public to use and improve. They've enabled collaboration across countries and cultures for the tremendously productivity-enhancing and to many enjoyable and inherently liberating projects such as Wikipedia, the GNU/Linux operating system, and Android. If copyright is to cover AI, it should also ensure the continued viability of projects that "promote the progress of science and the useful arts" by using their exclusive rights to ensure public benefit directly.

7.3. "Unlearning". Keeping an AI system from drawing a particular inference seems beyond copyright law, which generally covers particular expression rather than ideas. But there are computer programs to detect similarities in writings and pictures and some compare to large collections of previous term papers and commercial music. Comparing an AI system's output to databases of copyrighted input for extensive similarity (which the user could rewrite before ultimate use, or which the AI might do in its output process) might be feasible and a reasonable balance between the copyright system's coverage of expression and the broader public interest in access to concepts (inventions being subject to much shorter-term and harder-to-get patent, if at all.)

8.3. Case law. Please also consider Mattel v. MGA, 616 F.3d 904 (9th Cir. 2010), about "Bratz" dolls, discussing how copyright must often cover a given work only narrowly to avoid restricting practical access to general concepts and to expression in the public domain:

"*Mattel argues that the sculpt was entitled to broad protection because there are many ways one can depict an exaggerated human figure. It's true that there's a broad range of expression for bodies with exaggerated features: One could make a fashion doll with a large nose instead of a small one, or a potbelly instead of a narrow waist. But there's not a big market for fashion dolls that look like Patty and Selma Bouvier. Little girls buy fashion dolls with idealized proportions — which means slightly larger heads, eyes and lips; slightly smaller noses and waists; and slightly longer limbs than those that appear routinely in nature. But these features can be exaggerated only so much: Make the head too large or the waist too small and the doll becomes freakish, not idealized.*

The expression of an attractive young, female fashion doll with exaggerated proportions is thus highly constrained. Cf. Data East USA Inc. v. Epyx, Inc.*, 862 F.2d 204, 209 (9th Cir. 1988) ("Because of these constraints, karate is not susceptible of a wholly fanciful presentation."). Because of the narrow range of expression, the preliminary sculpt is entitled to only thin copyright protection against virtually identical copying. Cf.* Ets-Hokin v. Skyy Spirits Inc.*, 323 F.3d 763, 766 (9th Cir. 2003) (photo of vodka bottle merits only thin protection because of limited range of expression);* Satava*, 323 F.3d at 812 (similar).*"

Google v. Oracle and Mattel v. MGA support free access to aspects of a work where snippets of expression are essentially equivalent to the underlying, uncopyrightable concept. The precision of terms of art is especially important for nonfiction works and paramount for works that are to be interpreted by a computer, such as programming code, rather than used directly by a human. Because AI systems look for mechanical similarities in a way that does not rise to the level of human judgment, verbatim use of passages within the system (or interconnected AI systems) seems important to their internal functionality for any purpose, even if Warhol Foundation v. Goldsmith would indicate that the final product an AI is used to make should be more markedly different from the original if it's to be used without a license in order not to circumvent protection of the original.

8.5. Market impact and fair use. It seems reasonable to consider an AI system's *prima facie* infringement on particular copyrighted (and unconsenting) works in light of the overall market for such works: copyright's intended incentive to contribute to a robust market would be undermined by allowing all of the other works to pick off each one in turn. But, the market impact of AI's use with public-domain or licensed materials should count against the overall market downturn to determine how much of a work's overall competitive downturn is really due to its having only a little incremental value over an alternative that market participants choose to offer up. Moreover, the presence of numerous alternatives to a given work suggests that much in the field consists of uncopyrightable principles and concepts (see Mattel v. MGA.)

9 through 9.3; 27. Infringement and consequences. I'm inclined to prefer an opt-out approach because I think it would make AI more powerful, innovative, and useful faster, with tremendous direct and indirect benefits, but I don't create works that copyright may be important to monetize for a living.

More importantly, considering that AI apparently needs to draw on a large volume of input to function well (with breadth being exceptionally important in some kinds of research to ensure one reaches rare circumstances too), the possibility of very harsh penalties for some kinds of copyright infringement, and the possibility that source material that should be excluded would get inadvertently scooped up or even maliciously fed to a system, it's essential to the viability of this important technology that operators be able to correct any violations at a fair and generally modest cost.

Moreover, all of this is relevant only to the extent use of materials for AI training is not itself fair use. I think at least some kinds of scanning for concepts without broadly reproducing expression could be.

18, 20 and 21. Copyrightability of AI outputs. It seems clear that copyright could cover the internal program code and selection of input data and prompts to an AI system. Further, as Google v. Oracle appears to reflect, modern computer programs consist of multitudinous interconnected functions that no particular individual, even a programmer, is likely to know all the workings of. Many long-standing computer art practices, like an "impressionist" photo-editing function, create somewhat abstract and unpredictable effects. As do some long-standing lower-tech practices like Jackson Pollock's popcorn-ceiling-esque paintings. Would that make uncopyrightability of "AI art" a false dichotomy as to lack of

authorship, and an unfortunate one insofar as copyright may provide good incentives? Computer based art could contribute tremendous technical expertise from a variety of disciplines and help one try lots of great ideas *fast* in addition to saving money, reducing environmental impact, and making copying (even charged-for copying) of the result efficient. Although, art that draws heavily on fair use of myriad preexisting works might justify only a very "narrow" copyright of the particular result purposely or fortuitously struck.

28; 30-31. Labeling of AI creations; personal-likeness related rights. In addition to state-law rights to profit from publicity associated with oneself, AI may facilitate misattribution of statements or actions to a person which may defame them, relatedly cast them in a false light, impair their ability to speak credibly, or dilute their trademark-like goodwill. I think it's better that users look out for the occasional AI-related error or awkward take on reality rather than throw out a very valuable and playing-field-leveling system with this bathwater, but it might be reasonable to require large commercial operators to correct or flag such problems, especially upon notice, and more generally to alert users to AI-generated content. Voluntary best practices could be a good solution as AI businesses may want to be on good terms with authors, artists, and their supporters.

34. Any other issues. To the extent that copyright would cover using a work as an input to an AI system for output that would not be considered infringement if created by a human, please consider:

- Noncompete equivalents. Whether contracts of adhesion covering one's preexisting, current, or future works would tend to disadvantage "ordinary" employees and reduce innovation similarly to proliferation of non-compete clauses. The Administration proposes to restrict those, at https://www.ftc.gov/legal-library/browse/federal-register-notices/non-compete-clause-rulemaking .

- Antitrust concerns. Whether market and/or regulatory demand for answers that are as reliably complete as possible, and practical necessity to complement one's own efforts with AI, would make access to source material for AI systems a "natural monopoly" dominated by one business or a few with interlocking and interlocking rights.

- International competitiveness. Whether friction of broad copyright rules on the use of AI, or anti-worker or monopolistic practices that may enable, would make the US economy unattractive for innovation in AI and complementary industries. (California's culture of innovation has been attributed in part to banning noncompetes and China's rapid development has been attributed in part to lax IP enforcement.)

- Public rights in taxpayer funded works. Much discovery and art, as well as much relatively simple research of topics and practice of techniques that AI could draw aggregate value from, happens pursuant to government grants and at schools that are funded both directly by the government and through student loan subsidies. The White House Office of Science and Technology Policy apparently required that some publicly funded research be made available without charge. https://www.whitehouse.gov/wp-content/uploads/2022/08/08-2022-OSTP-Public-Access-Memo.pdf ; see also Bayh-Dole Act, Wikipedia, regarding patents on federally funded inventing.

/s/ Anand Desai
October 30, 2023