To whom it may concern at NIST,

My name is Karla Ortiz , I am professional concept artist, illustrator and fine artist.

I hope to submit to NIST concrete suggestions concerning generative AI guidelines, standards, best practices and even nomenclature for generative AI.

To address this, I suggest guidelines that focus on and enforce existing laws around Copyright, as well as potentially implementing new privacy preserving legislation that in a similar vein to the European GDPR. I would also suggest NIST, alongside the Copyright Office, continue to assert that copyrights require human authorship.

Furthermore, by allowing developers to openly train on Copyrighted material for all use-cases, the US and its copyright dependent industries are put at great risk of losing pivotal protections that not only foment innovation but also allows the creative sector, a sector that contributes 4.4% to the nation's GDP, according to the US Bureau of Economic Analysis.

To simplify this, there are a few core concepts that can help give us all a solid foundation to build upon.

Consent, Credit, Control, Compensation and Transparency.

These five pillars implemented as requirements for AI development promote safety for AI systems, industries and the general public. The principles of Consent, Credit, Control, Compensation and Transparency, pushes generative AI makers to be mindful of the data they utilize to train generative AI models. This would prevent the accidental infringement of authors and users rights, of violations of the general public's privacy rights, and a myriad of other issues that could have wide ranging effects in society.

NIST could offer solid and robust guidelines when it comes to the data utilized at any point of the Generative AI training process. In particular to prevent further exploitation of people's labour, privacy and intellectual property, no copyright, biometric or private data should be used to train GenAI without explicit consent, credit, control and fair compensation when applicable. In order for governments and citizens to be able to know if a company is following that first basic step of consent, NIST must push for the full transparency of model’s datasets, to be made available to industry and public scrutiny.

There should not be exemptions to copyright for AI training, AI companies must follow existing law like everyone else. Companies however can still build great models by the use of public domain data and expand via fair licensing to build models.

Another topic worth noting is that companies offering “Opt-out" to people whose data and work have been utilized without consent, is not a viable solution to the problem. Currently, techniques for genAI models to "unlearn" specific samples of data are extremely costly even for moderate or small models, and completely intractable for state of the art generative AI models. Due to this it is near impossible for algorithms to forget data once it's been trained upon, outside of destroying the algorithms and training sets and starting from scratch. Thus, while Machine "Unlearning" is still a nascent stage, current "opt-out" procedures are largely ineffective and unreliable.

Even if opt-out procedures were perfected, this still misses the mark. Opting out puts the onus on the public to police the use of their data. Given the multitudes of AI models that have proliferated in just one short year, asking a member of the public to constantly monitor the use of their own data by AI companies that are constantly multiplying seems to me to place the burden on the wrong party. For starters, it places an undue burden on people who may not be well versed with the technology, people who may not know the language, people who may have physical impediments, people who may have time limitations, or people who are not on the internet, are unaware their data is in the possession of technology companies and being used for commercial profit.

On a practical level opt-out makes no sense and does not address the issues at hand. The only answer to this dilemma is via an “Opt-In” only approach. Where companies in order to expand the quality and variety of their datasets approach rights holders to engage in good faith negotiations of fair compensation for potential use (with sensible limitations) of their data and works. This approach follows the earlier principle established of Consent, Credit, Control and (Fair) Compensation.

It is important to act measured but also with haste as the various economic harms of Gen AI are soon to be felt or in some cases already impacting everyday Americans. These harms include but are not limited to the violation of intellectual property rights, diminished jobs, lost jobs,deep fakes, non-consensual pornography , child abuse material, invasion of privacy rights, identiry theft, scams, potential to make copyrights obsolete and stifling innovation.

Additionally more research is urgently needed and NIST could set standards to study the impact to industries and immediately implement best practices to prevent job losses.

Nist could also close the research to commercial dataset loops holes, otherwise known as "data laundering", employed by many offending GenAI companies. This particular loophole is when a vast dataset is made via web scraping practices for research purposes, making full use of data exemptions that come from scientific research, to then that same dataset to immediately be used for commercial purposes, where that commercial entity would’ve had to follow different guidelines to acquire specific data, essentially “laundering” that data. This loophole must be closed to prevent future abuse of any future technology in the future.

On naming conventions, using just "AI" is too broad and gives the wrong impression to the public and society. Generative AI is good, Machine Learning is better. For the media, the term "AI Art" should never be used. Generative Ai Media, or regenerated content, or synthetic media are all better and clearer terms to use. Regardless better terms are needed to avoid further anthropomorphization of algorithmic models, a necessary step

for better development of guidelines and best practices .

Other suggestions for NIST include recommending Government to pass laws making it simpler for those who have had their data used to train Ai models without consent to vindicate their rights in federal court and seek statutory damages.

Other suggestions include potential taxing or other deterrents for companies that replace jobs with GenAI.

Nist could recommend Copyright registration to be sufficient for copyright infringement litigation as long as registration is gained before trial. Copyright registration could also be easier, and allow batch copyright registration for artists

Regardless, as NIST continues to forge guidelines for potential best practices, it must avoid the risks that potential regulatory capture brings. NIST can avoid this by ensuring stakeholders in affected industries have a voice and say in policy and regulatory processes.

Lastly, once again, as long as the five guiding principles of Consent, Credit, Control, Compensation and Transparency are followed, I believe we'll find the right balance for our technological and creative futures.

Thank you

Sincerely

Karla Ortiz