

1. These algorithmic generative tools, which I refuse to call AI on several principles (first of which that the term is just hype-based sector marketing for applied statistics and I don't buy the hype) pose a great risk for creatives and the general population without providing much tangible benefit to society.

Before the WGA strike resolved with contract-baked protections for writers, these algorithmic generative tools were a massive downwards pressure on wages and working conditions, and their permissible uses was one of the major points of contentions which the studios initially refusal to acknowledge concerns about forced the writers into striking for the sixth longest period in hollywood's history (148 days). The currently still on going SAG-AFTRA strike is contesting similar issues and it's likely that every other strike looming over the entertainment industry, if or when they happen will also hold these tools's permissible uses in contention as well.

These tools have been demonstratably bad for creatives, with several anecdotes from commission workers that independent commission work is down across the board as these tools are being adopted to replace the labor cost of hiring artists or writers; and their results are mediocre at best, most likely sorrid in the boring sense of the word, and dangerous at worst. In the information industry specifically, mass-produced and dangerously misinformative texts are being published, and the rate at which these can be made is putting a massive strain on our already stretched thin ability as a society to manage dangerous misinformation.

These tools weren't around during the 2016 election when the cambridge analytica scandal basically attempted to sell our free election to russian interests: and now that theyr here and untrue articles can be mass produced to spin false narative with minimal human involvement, I shudder to imagine what the 2024 election cycle might look like.

They are dangerous and their added value does not generally justify their dangerous nature; and that's before considering the ethics of data set selection.

4. While I do believe that international consistancy with regard to regulation is almost always helpful, I do not believe it is necessarily a major factor in this area. Training these machine learning algorithms is an time, energy, and resource intense process, and it'd likely be seen by corporations and individuals as massively wasteful to train multiple variations of the algorithim to comply with many different individual standards. Any regulation in one jurisdiction is likely to shape the overall world standard that corporations adhear to for broadest market appeal.

5. Yes, new legislation is necessary in this area and not only is it necessary: it is overdue. The legislation should cover how data sets are allowed to be sourced, who is allowed to receive dividends or benefits from a data set, means for providing remedy when a data set in found to be in violation of ethical or legal standards, and how algorithim owners can be held accountable for the contents of their data sets, as well as specific punishments for the malicious use of these tools by developers and/or end users.

6. With the exception of the music industry where machine learning generative algorithms have been developed using public domain and/or volentarily liscensed material as training data (primarily because the centralized powers of the music industry dissuaded anything napsteresque from happening), these algorithms are trained using data sets in the petabytes, and it's almost impossible to say exactly what they contain because the data set is considered proprietary. However, it has been demonstrated through specific prompts that language model algorithms like GPT-3 were trained using a data set which contain significant amounts of text scraped from the Organization of Transformative Works's Archive of Our Own (AO3) which exists to archive fanfiction, texts of transformative material created by (and therefore owned by) individuals who most often do not own the rights to the source material's copyright: AO3 is allowed to exist under fair use doctrine as a nonprofit: GPT-3's inclusion of these scraped texts in the data set with the permission of the texts' authors and without the permission of the texts' source materials' copyright holders demonstrates two layers of infringement upon intellectual property.

The majority of image generating algorithms are similarly infringing upon the works of artists whose art was scraped without permission or compensation, as

well as private medical records, and the intellectual property of large intellectual property owning corporations like Disney.

They were (almost) all built upon data sets of scraped material that were so large as to be prohibitively expensive to individually license, that the executives training these algorithms did not bother with the question of data set ethics nor copyright infringement.

7.1 It depends upon blackbox implementation and intended output how exactly the algorithm uses the images during its training.

With language models, they are typically trained by being provided a sequence of words; for example, "The quick brown fox jumps over the lazy ___" and made to guess the next word in the sequence. After guessing and being told whether it was correct or not, a lot of math happens, and the model gets better at guessing that the next word in that sequence should of course be "dog." This is then repeated a million more times across a million more sentences with various confoundations to prevent over-fitting until you get an algorithm which is very good at guessing, given any number of words, the next word in the sequence. The magic happens when you take the output of the algorithm and feed it back into it: it guesses the next word and the next word, and arrives at something resembling coherent thoughts but with no concern for truth or the meanings of the words. The only thing the applied statistics here creates is a string of words which are more or less very likely to follow each other.

With image generators, they are typically trained using an adversarial network, in which two networks compete against each other: one trying to create fake images and the other trying to guess which image being shown to it is fake. When the guesser guesses correctly, it points to the flaw it used to guess so that the faker can do a lot of math to fake better next time. and when the guesser guesses incorrectly, it does a lot of math to guess better next time. Typically, the faker generates its images by starting from random noise and through a series of transformations (as determined by the algorithm) fine-tunes that noise into something recognizable to a human viewer.

7.2 There aren't inferences in machine learning; Not really. It's all just applied statistics to determine the most likely output from input. The machine understands by the end that it will get a good score in guessing the correct word if it guesses "dog" but it does not understand what a dog is: only that the answer "dog" likely earn it a good score. Similarly, the image generator doesn't understand what it is drawing in a human way: it only understands that specific convolutions in the random noise seeding the image will be more or less likely to confound the guesser (and by extension a human viewer) but it doesn't understand what those convolutions create the illusionary image of: it does not see a dog. It sees a sequence of pixel transformations. Inference is entirely too generous a personification of these algorithms.

7.3 If it is possible for an AI model to 'unlearn' anything, like a specific piece of training data, the method of doing so is not currently known to our understanding of mathematics. It may be possible in the future: but, it is not possible in the present.

7.4 Yes. It has been demonstrated that even without access to the training data, it is possible to identify if a machine learning algorithm has been trained on a particular training material: These algorithms function based on likelihood determined by applying statistics to the training data, and it's impossible for them to generate patterns which they've not been exposed to in training. ChatGPT is able to replicate the word-order patterns of the ABO genre, therefore ChatGPT's training data contained works from which it observed those patterns and performed math on them. If the use of an artist's name in the prompt manages to generate images in that artist's style or images mirroring that artist's compositions, that artist's images must have been present in the model's training data for it to have correctly incorporated that pattern into its model.

8 None.

8.1 While an algorithm itself is a wholly different and technically transformative thing, the outputs said algorithm produces are not transformative but derivative: and because what is being sold is not the algorithm itself, but the algorithm's capacity to generate derivative outputs,

I believe that machine learning algorithms for the purpose of copyright should be considered derivative works not subject to fair use protection.

8.2 Data sets are very directly derivative works which wholly rely upon the works they are derived from to exist. No source material?: No data set. They never qualify for fair use in my opinion.

8.3 It shouldn't. It doesn't.

8.4 The developers of any technically impressive machine learning algorithm is using upwards of terrabytes or even petabytes of data to train their algorithms. The only difference the size of a given data set makes on the argument of fair use is whether any given copyright holder will be able to find where in the data set their infringed upon material is: It is a means of offuscating infringement via sheer scale of the needlestack.

8.5 As stated in my answer to question one, artistic commissions and independant contract work are, admittedly anecdotally down across the board, coinciding with the rise in availability of imagine generating algorithms: they are being used and marketed as products which cheaply replace the expensive labor of trained artists. Writers were before the strike yielded them in-contract protections against machine learning facing the prospect of replacement and depressed wages of machine generated scripts being mislabeled as original source material.

9 Affirmatively consent.

9.1 All use of copyrighted works in training data.

9.2 I do not believe opt out is the correct approach: some technical flags are currently being pioneered by websites, but the individuals leading the machine learning algorithm industry have demonstrated a lack of care for ethics and would likely not be deterred by a NoAI metatag.

9.3 If it is not feasible to get consent from all potential copyright owners, the project's data set should be scaled back or its staff increased until it is feasible. If this results in worse algorithms over-all or lower profit margins for executives, corporations, shareholders, etc, so be it. Ethics are more important than profit in this regard.

9.4 Because there is currently no mathematical way to untrain a piece of data, any failure to comply with data set regulations should be considered an immediately poison pill for the model trained on that data set and all subsequent models refined downstream of a tainted model. If this results in lower profit margins for executives, corporations, shareholders, etc, so be it. Compliance to regulations is more important to profit in this regard.

9.5 It is my opinion that consent should be required by both the copyright holder and the individual or team of creatives who made the copyrighted work. While I consider this the ideal system, I do not know how such a system would be implemented to have work under current copyright law.

10 Through individual negotiations or through negotiation with a collective representative body.

10.1 Whether direct voluntary licensing is feasible is, in my opinion, immaterial to whether it should be the default.

10.2 The most well-suited organization that currently exists to collective voluntary licensing are the organizations which manage the creative commons license.

10.3 No.

11 The responsibility of licensing should belong to the data set curator. Developers should have required due diligence, checking that all appropriate

licenses were obtained by the curator before incorporating the data set into their training data. In the case of failure of the curator, they should be liable for damages to both the infringed upon party and any developer who used their data set. In the cases of failure of the developer to perform their due diligence they should be liable for damages to the infringed upon party.

12 No. While it is possible to determine that a particular piece influenced the algorithm during training, it is not possible using current mathematics to quantify how much of an influence sed piece had on a particular output.

13 It would likely make them much more expensive to develop. This is a good thing as they are currently only so cheap through the massive, almost industrial, infringement and exploitation of vulnerable people.

15 Yes. 100% transparent and independently auditable records of training data should be manditatory.

15.1 They should be required to keep a copy of their full data set, including metadata and documenting the sources of outsourced data sets.

15.2 A method to access and sortition the record of their data set should be made available anywhere where the model trained on that data set is available for use or download.

15.3 Any data set incorporated from a third party must be fully documented as if it'd been curated by the developer themselves.

15.4 Hardware costs for data storage and server costs for access to the data. The cost would depend upon the size of the training data and the popularity of independant audits on the data. I would love to be able to trust companies in this sector but quiet frankly they've already shit the bed and we're past that trust thermocline. If they want to play ball they gotta do so naked.

16. An obligation of due diligence: 'We purchased this data set from this curator which has recorded you as having consented to your inclusion in the data set: if that is not the case, please inform us and the curator to clarify any miscommunication.' This protects both the copyright holder and the developer from a failure on the data set curator's part to secure the correct liscense for material in their data set.

18 No.

19 No.

20 No. No. Yes.

21 No, it does not. No, it would not. If no one could be bothered to create it, why should we bother protecting it.

23 Yes.

25 It will depend on a case-by-case basis, but generally the developer of the model, the curator of the data set that introduced the infringed upon work would be liable condition on proving failure to secure the proper liscense, while an end user would be liable conditional on proving intentional infringement (eg likely use of infringed upon partys name or IP in the prompt).

25.1 I do not believe they do because I believe that all models should be compelled to be open source, at least in their training data.

28 Sometimes, depending on the outputs capacity to decieve or propegate misinformation. Any output which a reasonable person would at first glance perceive to reflect an authentic facet of our reality rather than be an algorithmically generated output should require a label. For images, that

means photorealistic images generated. For text, that means words styled after news articles, interpersonal communication, scientific papers and other academia, etc. The more likely it might be mistake as not the output of an algorithm and the more detrimental such a mistake would be, the more strongly the requirement of a watermark or label should apply to it.

28.1 The developer of the algorithm is responsible for making sure outputs comply with regulations: the publisher of outputs are responsible for making sure outputs are not sanitized such that they no longer comply with regulation.

28.2 None that haven't been overcome by regulations for media before this.

28.3 It should be proportionate to the potential harm failure to comply in that case caused, could have caused, or was intended to cause.

29 For most cases, the tools that exist or are in development to identify algorithmically generated outputs are themselves machine learning algorithms. They are subject to the same trapping and pitfalls that the algorithms they're seeking to combat are, as well as the additional problem that they by the implementational necessity of needing to generate training data for them to be trained to detect lag behind their opponents by a significant margin. This is an area that needs extensive research into forward progress.

32 Yes. I do not know implementation details, but yes.

34 Because of a machine learning algorithm increasing the capacity for harm caused by an individual using the tool, cases where an individual slanders or defames another person, spreads harmful misinformation, or infringes upon copyright should be considered separately in a court of law based on whether a generative algorithm was used as a tool in enabling the offensive action; similar to how "assault" and "assault with a deadly weapon" are in many jurisdictions considered legally distinct.