John LaClaire (individual comment)

Copyright and Patent are an explicit bargain between society and creators – a grant of exclusive rights, for a limited time, in an authors work serves as incentive to create new works and to make those works available in the marketplace and thereby "promote the Progress of Science and useful Arts".

This bargain is fundamentally progressive – during the limited period of exclusive rights others can learn from works shared in the marketplace; upon expiry of that period society at large gains the benefits of the creator's contribution to 'progress' of science and useful arts.

A definitional challenge in granting copyright or patent rights is that 'progress', and its underlying elements, are ambiguous. Both originality and creativity are problematic to define, yet giving an author the benefit of exclusivity in the absence of "Progress of Science and useful Arts" upsets the bargain upon which the rights are founded.

From: https://supreme.justia.com/cases/federal/us/499/340/

> The most fundamental axiom of copyright law is that "[n]o author may copyright his ideas or the facts he narrates." *Harper & Row, Publishers, Inc. v. Nation Enterprises,* 471 U. S. 539, 471 U. S. 556 (1985).

To distinguish a copyrightable work from mere recitation of 'facts', the supreme court resorts to declaring a "modicum of creativity" is required. Creativity and originality are usually clear cut for lengthy works, but as works are carved up into small pieces, as authors seek protection for phrases of text, music or individual lines of software code, the task of determining originality begins to resemble the task of determining novelty in patent law.

Turning to Generative AI, an open question is whether the output of models trained on particular sets of facts possess a modicum of creativity, or are merely a recitation of facts? Are the algorithms used by generative AI creative, or do they merely emit an algorithmic rearrangement of facts?

If we take one step back and focus on progress and utility, rather than creativity, we can ask a slightly better question. Do the outputs of generative AI "promote the Progress of Science and useful Arts"? The utility of generative AI itself is certain – just as all modern computers can do math more quickly than the fastest human, generative AI can get some types of results faster than a person could produce equivalent results. Generative AI systems, individually, and as a class of algorithms, clearly have utility.

The result (outputs) of a particular generative system, however, remain algorithmic – that the algorithm is very complex does not change the fact that the result is algorithmically derived from facts (the input prompt and training set). In most [all?] cases the algorithm (model) used by Generative AI software has been trained to emit output that represents a "most probable" or "most representative" output based on the prompts, training set, and algorithm used to train the "model". A wholly algorithmic output seems to be the antithesis of creativity or novelty – the arrangement of the output is surely less obvious than an alphabetical ordering – yet it is an algorithmic arrangement of facts, differing only in the complexity of the mathematics used to "order the facts".

That a “most probable path” from a complex data set can now be calculated offers the possibility that creativity and/or novelty could be objectively defined and potentially even quantified as a variance from that “probable path”. An objective and quantifiable definition of creativity (or novelty) would protect society from unjustified issuance of intellectual property rights where expressions or ideas are algorithmically probable, rather than truly creative or novel. This stands in contrast to those expressions or ideas that go off, or beyond, the most probable path, wile also having utility. It is these unobvious ideas or expressions that truly advance science and the useful arts.

Tools that better distinguish between true advances and mundane or probable ideas and expressions would reduce the grant of exclusive rights on ideas or expressions that are not material advances, while rewarding genuine advancements. By better sorting the “what from the chaff” possession of copyright or patent rights becomes more valuable and more objectively defensible.

A tool or method to objectively measure creativity or novelty might involve running a work submitted for copyright (or patent) essentially ‘backward” though a generative AI model and computing how unlikely it would be for that model to generate that particular output. Another possibility is to add the submitted work to the training set and measure how much impact the work has on the model, though it is the impact of the change, not the quantity of weight changes, that would need to be measured.

Another great challenge in creating such a tool is that the result will, of course, depend heavily on the training set – so the issue of training set bias would need to be better understood and addressed, and of course the model would need to “learn incessantly’ to keep up with progress. I lack the expertise to go further than the above naive comments and proposals, but I would encourage the copyright office to promote research into use of Generative AI to objectively define and quantify “novelty” or “creativity”.

If Generative AI does indeed offer an opportunity to quantify “advances” in science and the useful arts it would put intellectual property law on far more solid and objective foundations. If it is possible at all, getting there will not be easy. Use of any data set smaller than the sum of human knowledge would necessarily imply that some extant works would be judged to be novel. Getting to a sufficiently unbiased model that never the less constitutes the sum of human knowledge is no small undertaking. Getting close enough to have utility in assisting human examiners in assessing the novelty of ideas submitted for patent or the creativity of works (or small portions of works) may, however, be a tractable problem.