

Komolgorov vs. TESCREAL: Can complexity theory offer a bright line within which the use of LLMs is ethical?

Jonathan A. Poritz

jonathan@poritz.net
poritz.net/jonathan



23 June 2025

Association for Learning Technology

OER25: Speaking Truth to Power:

Open Education and AI in the Age of Populism



This slide deck, except where otherwise indicated, is by [Jonathan Poritz](#) and is released under a [Creative Commons Attribution-ShareAlike 4.0 International License](#). This version: 23 Jun 2025 14:08BST
The slides, also in editable form, are always available at poritz.net/jonathan/share/ALTOERKvT/.

Positionality acknowledgement

From no benefit of my own, I benefitted from a WEIRD background and, largely, present.

This **WEIRD** stands for:

- W**hite [or **W**estern]
- E**ducated
- I**ndustrialized
- R**ich
- D**emocratic

The **W** in WEIRD



¹“Silence = Genocide with Palestinian flag by Daynarama is released under [CC0 1.0](https://creativecommons.org/licenses/by/4.0/).

The **D** in WEIRD .. and the *F-bomb*

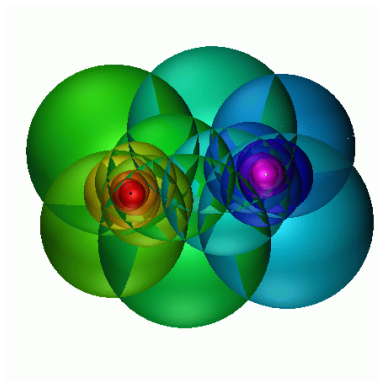
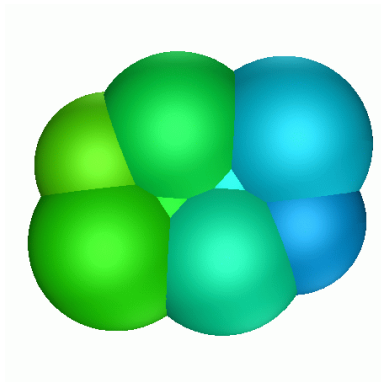


Democratic is questionable in many places at the moment.

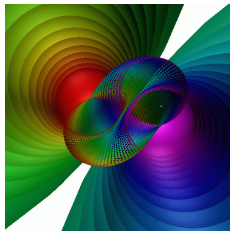
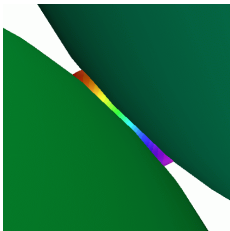
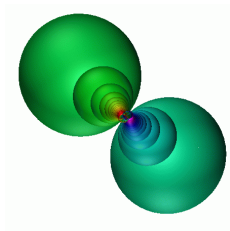
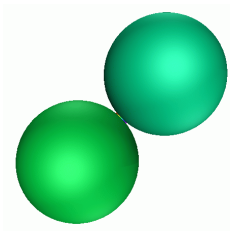
Sometimes I think the only thing we should be talking about is how to prevent the [further] rise of **fascism**.

²This image, by [Jonathan Poritz](#), incorporates the image [Fascies](#) of Višeslav with license [CC BY-SA 4.0](#)) from Wikimedia Commons, as well as an image with license [CC BY 4.0](#) created by Poritz adapting the [Global OER Logo](#) of [Jonathas Mello](#), which also has license [CC BY 4.0](#).

Disciplinary positionality: math 1



Disciplinary positionality: math 2



Secure intrusion-tolerant replication on the Internet (with Christian Cachin) had a bump in success after blockchains came along.

Universal Gates in Other Universes, on complexity theory in universes with different physical laws, is the paper of which I'm proudest.

FLOSS v capitalism

[FLOSS = **Free/Libre/Open-Source Software**].

Even in IBM – one of the forces of evil from when I was a young geek! – FLOSS absolutely rules. As it does across the Internet, in a way that has gone completely under the radar of the modern world.

This is simply because FLOSS is better, e.g., because

- *with enough eyes all bugs are shallow* [transparency!], and
- FLOSS almost always has *many* more features than corresponding commercial software [diversity increases creativity!]

But **not** because we were suddenly not bad actors – as their purely *this dinosaur is dead, but we can bring in a pile of money before it realizes that and falls over* approach to DRM showed, as I could see from within the belly of the beast.

In companies of all sizes, I had direct experience of how the capitalist feedback mechanism lead instead to immoral *profits over people*, except when constrained by policies and regulations.

Truth-seeking exercises, or the *good scholarship*TM

Policy and law/regulation can only provide a effective humane counterbalance to (rapacious late-state) capitalism if they are operating with detailed, insightful, and accurate information.

Hence governments go through *truth-seeking exercises* ... which we scholars recognize as nothing more (or less) than *good scholarship*TM (or *the scientific method*).

Some crucial elements of *good scholarship*TM include:

- **impartiality/disinterestedness**
- **transparency**
- **diversity**
- **ethics** and even
- **extraordinary claims require extraordinary evidence.**

Emily Bender and Alex Hanna of [Mystery AI Hype Theater 3000](#) say “Always check the footnotes,” but I think we scholars have to go even further and **always read corporate pronouncements as if we were doing peer review!**

Impartiality

Impartiality is important because of bad actors: think of examples from Cicero's cry of *Cui bono?*

all the way through the MD, paid by the tobacco industry, who said that cancer causes smoking, not the other way around.

...And also perhaps even because of self-deluding good actors, as in the wonderful quote:

It is difficult to get a [person] to understand something, when [their] salary depends upon [their] not understanding it!"

(from **I, Candidate for Governor: And How I Got Licked**, by Upton Sinclair, University of California Press, 1934; edited to be gender-neutral.)

I'm sorry to say that a lack of impartiality disqualifies a huge proportion of the academic computer science community from being reliable experts, since there has been a vast take-over of this discipline by wealthy computer industry influence – the [Coalition for Independent Technology Research](#) has assembled interesting research on this topic, and is doing more.

Transparency shouldn't be hard to sell to the open community.

But note that while we struggle to push for more openness in education and scholarship, it has been obvious for a long time that it is simply a basic requirement for doing good science. For example, transparency is a *sine qua non* for good science, via such things as pre-registering hypotheses and rigorously published **all** data collected in an experiment.

These techniques of good science are important to prevent bad actors from doing **p-hacking**, such as for example when a bad actor with deep enough pockets simply does an experiment twenty times until they get a result which goes the way they want, regardless of the other runs' results, at a significance level of $p = .05$, which is considered (even by the courts in the US!) to be the standard significance level.

Diversity is under attack in the US and, yes, in Europe, right now.

I've already mentioned that the diversity of voices in FLOSS projects is one of the reasons they tend to have so many features and to be so robust, flexible, and powerful.

There is also evidence in the sociological literature that diversity can make for creative teams, which I'm sure we've all seen.

I would say that the scholarly community has come to understand ever more clearly – and it is a long, even unending process to make progress here – that **ethics** must be a core aspect of the work we do.

Examples like

- the Milgram *obedience to authority* experiment,
- the Zimbardo *Stanford prison experiment*,
- the *Tuskegee Syphilis Study*,
- the *contraceptive trials in Puerto Rico* on women without their knowledge or consent, and
- too many others to name, alas.

In fact, I would argue that the basic principle that

Institutions of higher education are conducted for the common good and not to further the interest of either the individual teacher or the institution as a whole.

from the [American Association of University Professors' 1940 Statement of Principles on Academic Freedom and Tenure](#)³ squarely places **the good**, and therefore ethics, at the center of what we do.

³ which was built on the [1915 Declaration of Principles on Academic Freedom and Academic Tenure](#)

Extraordinary claims

That **extraordinary claims require extraordinary evidence** was described in this catchy way by the astronomer and science communicator Carl Sagan.

But the idea behind this quote goes back to Occam, Laplace, and Hume, and is a fundamental tenet of empiricism.

When the extraordinary claim is which will entail enormous expense or have an impact of many lives, it seems like this principle is not only good philosophy of science, but also grounded in prudent common sense.

For example, before we rebuild our curricula for wide swaths of education, I would like to see some really irrefutable evidence that a newly proposed tool is going to transform for the better our students' ways of learning and post-graduate lives.

One last point which is perhaps not as widely accepted or quoted – maybe simply because it is so obvious? – is that it is better to reason about something **bare-brained**, without analogy or metaphor, if it is possible to do so.

There is been interesting – even thought-provoking and inspiring – work done on metaphors for AI and other edtech by Nate Angell, Yasser Atef, Maha Bali, Anuj Gupta, Leon Furze, Anna Mills, Marin Weller, Audrey Watters, and others.

While I admire this work, I cannot help but notice that there is an undercurrent in discussions of AI of how sophisticated and mysterious the tech is, how mere mortals should not hope to wrap their heads around the details, certainly not bare-brained and without a protective layer of metaphor. Since this puts the very non-disinterested originators of the tech in the privileged position of oracles who, by controlling the metaphors, can control the discussion, it makes me very concerned.

TESCREAL digression

Timnit Gebru and Emile P. Torres published in [The TESCREAL bundle: Eugenics and the promise of utopia through artificial general intelligence](#) some very important research on the strange set of beliefs which seem to have almost a stranglehold on the leaders of the AI industry. Here TESCREAL stands for:

***T**ranshumanism, **E**xtrapianism, **S**ingularitarianism, **C**osmism, **R**ationalism, **R**effective **A**ltruism, and **L**ongtermism.*

These all seem like the wild-eyed 2AM rantings of a stoned high-school student who has read too much science fiction and understood none of it .. and yet because of these belief systems, we have immensely wealthy and powerful people stating that it is perfectly reasonable to accelerate climate destruction if it hastens the arrival of some imagined future in which trillions of uploaded minds can, a few years sooner, be enjoying “life” in their cyberutopia.

Some *good scholarship*TM steps

I would feel a little guilty about merely hectoring you for the academic community's failure to live up to *good scholarship*TM. So let's do two new things in this last part: Talk about what AI really **is** – not “what AI *is like*,” but what it actually is.

This is weirdly easy....

What is AI? (part 1)

The recent technical innovation which is used to justify the hype is based on something called a **Large Language Model [LLM]** or **Generative Pre-trained Transformer [GPT]**.

LLMs consist of a data structure called a “neural network”⁴ and two accompanying algorithms:

- one which processes (hundreds of) TB (or more) of *training data* - text and data usually copied from the internet (“scraped”) - to create (hundreds of) MB of values in the data structure (“model parameters”);
- another which uses the data structure to transform sequences of *input tokens* into sequences of *output tokens* in a way that maximizes the probability of the output given the input and the training data.

A “token” is a letter or word or image pixel, etc. Note the input tokens might be words in one language and the output words in another language, for machine translation, or outputs might be pixels, for image generation.

⁴ in a very poor analogy with biology

Aside on the Neural Network Data Structure/Algorithm

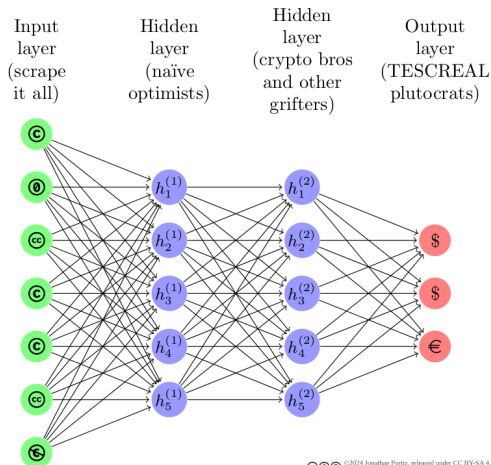
The neural network data structure looks something like this:

Input values are imagined to be put in each of the input layer nodes (those disks on the left).

Arrows show where values are used in simple formulæ to create values in each of the hidden layer nodes

The formal “output” of the model is a label (maybe a token) associated to whichever of the output layer nodes has the largest calculated value.

Real-World Neural Network Architecture



⁵ Real-World Neural Network Architecture is by Jonathan A. Poritz

Aside on *Attention is all you need*

This data structure has been around for a while, with some success in particular applications, but no one thought it was about to escape its creators and turn us all into paperclips.

Then a computer science paper called [Attention is All You Need](#) appeared in 2017, describing a particular way to use the data structure, together with some small additions that have been dubbed the "attention mechanism" (in a very poor analogy with "paying attention" as human minds do) in a particular way that allows relatively efficient scaling to extremely large versions of the data structure, trained on enormous datasets.

A "neural network" with the attention mechanism architecture is called a **transformer**, and when it is populated with parameter values created by processing some training corpus, it is said to be **pre-trained**: hence, GPTs.

Despite more than 185,000 citations in the scientific literature, the *Attention is all you need* paper does not present a deep new idea, or new framing of a deep old idea. Rather, it is a proposal for an incremental engineering change in how neural networks are used in applications which seem to produce very pleasing sequences of output tokens.

So:

$$NNs + \langle \text{attention} \rangle + \langle \text{huge corpora of training data} \rangle$$

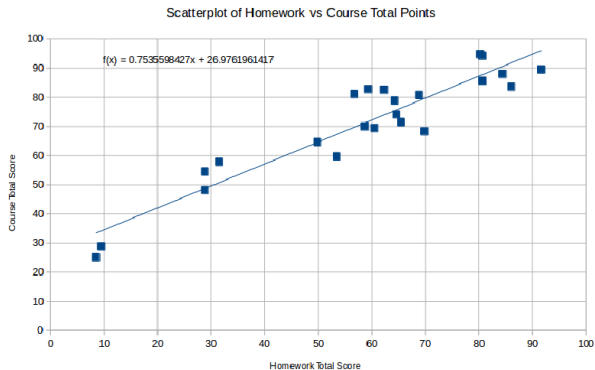
~>

algorithm that produces pleasing, statistically likely sequences of output tokens

In other words, a statistical model

Processing sample data to create parameters from which one can calculate output for new inputs is exactly what we call a *statistical model*.

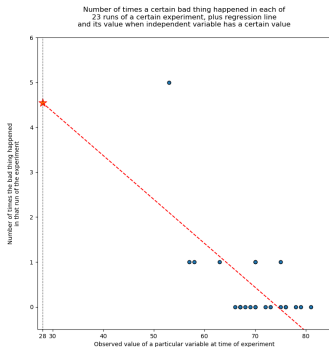
For example, you may have seen a statistical model called a “regression line.”



This *scatterplot* comes from an OER I wrote a few years ago for a stats class at my former institution, Colorado State University Pueblo. One calculates the coefficients of the equation of the line from the sample data and with this equation one can compute output values (“ $y = f(x)$ ”) from new input values (“ x ”).

Statistical models can do some wonderful things

Don't get me wrong, I love statistical models. They are very powerful. E.g.,



this one

might have prevented



Statistical models are not Skynet, nor do they have agency

Every sentence that uses “AI” as the subject of a sentence with an active verb is *prima facie* ridiculous:

- When Google’s “AI”-created search-results summary suggested to include a quarter of a cup of glue in a pizza recipe to prevent the cheese from slipping off when you pick up a slice, it was not “hallucinating.” Instead, it was outputting a statistically likely collection of tokens (words) in the context of its training data and the givens “pizza recipe” and “prevent slipping,”
- An “AI” will never “start a nuclear war”: a statistical model has no agency, the very idea is ridiculous. A bad human actor who used hacking techniques and nuclear codes extracted by a clever choice of prompt from an LLM which was trained with a little too much information might launch a nuclear weapon, but that’s pretty far from Skynet.
- An LLM cannot “summarize a scholarly article”: statistical models do not understand scholarly articles, that would be like saying that boats are good at swimming ... it’s a context error. Although an LLM *might* assemble sequences of words which are statistically likely in the context of the sequence of words in the article, and that could be a useful output in particular situations.

Compression uses statistics, too

When you compress a file, with zip or a similar program, you are making a statistical model of the file. For example, if a large file is 99% “X”s, and just a few of the other letters sprinkled in, you could make a compressed version of the file where every time you see the eight bits usually used to encode “X”, 1011000, you replace it with a single bit 0, saving 7 bits, and every time you see any other sequence of eight bits, representing any other symbol, you add a 1 in front, using an extra bit.

Exercise: describe how you would *decompress* a file produced by the above compression scheme: that is, how would you get back the full-length, original file from its shorter, compressed version? [Due by the end of the day; please don't use ChatGPT!]

In computer science, a **lossy compression scheme** is one which saves space, based on the statistics of the original file, but which when you decompress the file, doesn't give you back exactly the original file.

For example, if we modified the above scheme to store “X”s as 0 and everything else as 1s, then decompressing by replacing the 0s with “X”s and the 1s with “O”s, it would be lossy, because the decompressed file would have lost some of the original information: all of the original characters of file which were not “X” or “O” would be transformed into “O”s, while everything else would be recovered correctly. Depending upon how many of those mistakes there were, and what your purpose was for that file, this might be perfectly acceptable, or might not.

What is AI? (part 2)

Hence, another thing AI actually is, is a lossy compression scheme. It's huge pile of parameters are the compressed version of its even more huge training dataset, and choosing a prompt tinkers with the lossy decompression program, making it decompress in ways statistically correlated with the prompt.

Notice there is no creativity here, there is only lossy compression and decompression. Things that “look good” in a lossy decompression look that way because they look substantially like either the prompt or like something in training dataset, by dumb luck not too terribly mucked up from the compression/decompression losses.

What that means is that anything that looks good is a fortuitous assembly of fragments of human creativity in the training dataset. IAmNotALawyer, so I'm going to call that **theft**, the unacknowledged and unremunerated taking of someone else's work without permission. Even if not *legal* theft, this is certainly *academic* theft.

Maybe I am an attribution fundamentalist, but I say this theft makes all uses of LLM output presumptively academic misconduct. Hence, here is another statement of what AI is: a lossy compression scheme which assembles its outputs from stolen human creativity.

Kolmogorov to the rescue?

Except for one situation: suppose you are using an LLM to translate the prompt into another language, or to read a text out loud, or to generate a description of a complex scene. In those cases, the model output has roughly the same amount of information as the prompt, so the model is not (illegitimately, in my view!) using the creativity of other humans as found in the training dataset to pad its output.

It turns out there is a way to measure the information content in a file in a way that, I think, captures what we are trying to preserve, called the **Kolmogorov complexity** [**K-complexity**] of a file and defined to be *the length of the shortest program which, when run, produces that file as output*.

Proposal: Any use of an LLM is presumptively illegitimate if the K-complexity of the model output is larger than the K-complexity of the prompt; if the output instead is no more K-complex than the prompt, then it is possible, but not certain, that the use is legitimate.

I like this because perhaps it preserves some a11y-oriented uses of LLMs. There is still the question of whether all the other downsides of AI make even those a11y uses suspect ... and I would a million times prefer we spend more to achieve the same a11y outcomes in other ways, so I'm going to say that probably all uses of LLMs are illegitimate.

Conclusion

The “problem of AI” in the academy is not one of a new iteration of the Tragedy of the Commons, nor is it (only or primarily) another dystopia brought to us by late-stage, neoliberal capitalism, nor is it even a hyperscale version of the *enshittification* which people so like to cite these days.

The problem is that, as a community, we have been distracted by the metaphors and fables of autonomous, non-living intelligence as gods and demons that have flourished in many cultures for decades, centuries, and even millenia. We have been cowed by the immense sums of money being flung about by VC firms and tech plutocrats who have the ears (and, sometimes, one wonders, the ends of the leashes) of world leaders. We have been wowed by an uncanny mountain of weaponized pareidolia to imagine some rough beast, its hour come round at last, Slouches towards Silicon Valley to be born.

The antidote to all of this is our old tried and true friend, *good scholarship*TM, which we do well, and which we know how to apply to the most exotic of situations, if only we screw our courage to the sticking place.

I believe that the first organization, department, campus, country, continent, or even individual or small group of individuals who effectively applies *good scholarship*TM to resist AI effectively will quickly become a leader and will reap great rewards.

Contact Info

Email: jonathan@poritz.net

Fediverse: [@poritzj@mastodon.social](https://mastodon.social/@poritzj)

These slides:

poritz.net/j/share/ALTOERKvT.pdf

and all files for remix:

poritz.net/j/share/ALTOERKvT/

Or else use 

to go to my homepage, where you can follow
links to all of my shared materials.



<https://poritz.net/jonathan>