Publishing an Untested Design:

A Debate over Two Essays on Corpus Recombination and Obligated Composition

Flyxion
Independent Researcher

October 2026

Abstract

Two recent essays describe a system for drawing sources from a large corpus, proposing an argument from them, and developing the argument under recorded obligations. The first counts the combinations the corpus contains and specifies the generator that samples them. The second specifies the surrounding pipeline of classification, retrieval, role-based prompting and obligated composition. Neither reports a measured result, and the code both describe has not been exercised against a real model in the environment where it was written. This essay sets out the case for publishing them now and the case for holding them until some evidence exists, lets each position answer the other, and then identifies the conditions under which each would carry. It concludes that the decision turns less on whether the essays are premature than on whether their status statements are strong enough to prevent a reader from mistaking a design for a result, and it lists what would have to change for publication to be defensible. The essay takes no position on the merits of the underlying system.

1 The Documents and the Question

The first essay observes that a corpus of several hundred works contains tens of millions of unordered pairs and triples, specifies a short program that samples them, works through one drawn pairing, and argues that the large count measures the corpus's capacity for juxtaposition and not its capacity to yield essays. It carries two appendices, one specifying the program and one reproducing its source. The second essay specifies the pipeline around that program: four organizations of the corpus, retrieval of evidence, successive role prompts to one local model, a coordinating decision on readiness, and composition accepted by a ledger of obligations and not by a score. It also records defects in earlier code from which the design was recovered. A third document, an earlier version of the second essay, has been superseded by it, and the question of publication does not concern that document.

The question is whether to place the two essays in the public repository now, in the author's usual manner of publishing work as it is produced, or to hold them until the system has been run on a small sample and its first outputs have been read. The debate below is conducted by

two positions and not by two people. Each is stated in the strongest form its defenders could give, and neither is the author's.

2 The Case for Publishing Now

The first argument for publication is that the essays are specifications and say so. A specification of a system that does not yet exist is a recognized kind of document, and its value lies in fixing the intended behavior precisely enough that the later implementation can be checked against it. Both essays state in their abstracts that they report no demonstrated performance, and the first states that the rate at which drawn combinations yield an argument is the quantity the procedure exists to measure. An essay that disclaims a result cannot be accused of overclaiming one.

The second argument concerns the timestamp. The repository functions as a version-controlled record of when each idea was stated and in what form. Holding the essays until the system has been run would place the specification after the evidence in the record, and it would obscure the fact that the specification came first and the evidence was then gathered against it. This ordering is exactly what a pre-specified study requires, and the first essay proposes such a study. Publishing the specification before running the study is the only way for the public record to show that the study was pre-specified.

The third argument is practical. The author expects to improve most of the published essays later, either with more computing resources or through contributions from others who improve different parts. A specification that is public can attract the corrections that a private one cannot. The defects the second essay lists in the earlier code are specific enough that a reader could repair them, and the repairs would then be the first real test of the design. Holding the essays removes that possibility and gives nothing in return except the avoidance of embarrassment, which is a weak reason in a practice that already accepts publishing unfinished work.

The fourth argument is that the essays are comparatively careful about their own limits. They separate a combination from a brief and a brief from an essay, they decline to estimate yields, they mark the process vocabulary as provisional, and they distinguish detecting a citation from verifying it. If any unfinished work is publishable, work whose main commitment is to distinguish designs from results is a strong candidate.

3 The Case for Holding

The first argument for holding is that a design published with a count in its title invites a reading it does not intend. The title of the first essay contains a number of the order of nineteen million. The essay argues carefully that the number is the size of a space and not of a harvest, but a title is read before an argument, and the number will travel without the argument. A reader who sees the title and not the text may take the figure as a claim about output. The exactness of the figure under a provisional grouping rule compounds the problem, since an

exact number looks like a measurement.

The second argument is that the evidence for the central object is thinner than the prose suggests. The one worked example is described in its own record as an editorial example developed from inspected passages, not a recorded model response, and the generator's model call has not been exercised in the environment where it was built. The essays say this, but they say it in the middle of a long text, and the strongest sentences in each essay concern what the design will do. Readers who have seen many system descriptions know that the description usually fares better than the system, and a published specification borrows credibility from the form even when it disclaims the content.

The third argument concerns the second essay's treatment of the earlier code. It lists defects in scripts produced in an earlier exchange, and its claims about what those scripts do are checkable. They are also claims about material most readers cannot see, since the exchange is not published, and a published statement that a script contains an unsupported recursive pattern or a conflicting use of standard input asks the reader to trust the author's reading. That trust is better earned if the repaired scripts are published with the essay and shown to run.

The fourth argument is that two essays on one system, written within a short time and overlapping substantially, may be better presented as one after the system has been exercised. The second essay already absorbs an earlier version of itself. Publishing in this state risks a trail of superseded documents, each of which a reader may cite.

The fifth argument is about the loop the essays warn against. The corpus the generator samples is the repository itself, and essays about the generator will become eligible sources for it once published. The essays specify a rule excluding generated briefs, but they do not address the case of an authored essay that describes the generator, which can then be drawn into combinations whose subject is the generator. Nothing in this is prohibited, and the resulting self-reference is a reason to wait until the rule is stated and the first runs are logged.

4 Replies

The defenders of publication answer the first holding argument by observing that the remedy for a misleading title is a better title and a status line, not delay. A subtitle or an opening sentence stating that the number is a count of combinations and not a count of essays costs nothing and removes the objection. They answer the second by pointing out that the thinness of the evidence is the thing the essays are about, and that a reader who finds the system described with exactly this thinness has been told the truth. They answer the third by noting that the defects can be demonstrated in a few lines, as can be done for the recursive pattern, and that the demonstration can be included. They answer the fourth by observing that superseded documents are the normal condition of a version-controlled repository, and the fifth by agreeing that the self-reference should be stated and denying that it is a reason to wait, since the point of a pre-specified rule is that it exists before the data do.

The defenders of holding answer these replies by granting that a better title and status line would address the first argument, and by observing that this concedes the point: the essays

as they stand are not ready, and the question is only how little must change. They answer the reply on thin evidence by distinguishing a description that discloses its limits from a description that has not yet encountered them. The first real run will probably expose problems that no specification anticipated, and a specification published before that run may have to be withdrawn or contradicted in public. They answer the timestamp argument by noting that a pre-specification is only as useful as its fidelity to what is later done, and that a loosely held specification is easily revised afterward without anyone noticing. A hash or a tagged commit would preserve the order without requiring wide publication.

5 What Would Settle It

The positions agree on more than the debate suggests. Both accept that the essays disclaim results, that the disclaimers are accurate, and that a reader could mistake the design for a result unless the status is conspicuous. They differ on whether the status can be made conspicuous enough by editing, or only by running the system.

Three changes would put publication on firm ground without waiting for a study. The first is that each essay state its status in its first lines, in plain terms: a design specification, with no run of the full pipeline, and with the model call untested at the time of writing. The second is that the title of the first essay stop carrying the figure without the qualification the essay supplies, or that the subtitle supply it. The third is that the demonstrable claims about the earlier code be accompanied by their demonstrations, which for the recursive pattern is a single line of Python and its error.

Two further changes would strengthen the case for publishing but are not necessary for it. One is a small run: the generator, with its model call actually exercised, on a handful of draws, with the outputs published as outputs and not as findings. A small run does not measure a yield, and it should not be described as doing so, but it converts the statement that the call is untested into a statement that it ran and what it produced. The other is a rule, recorded before publication, for how authored essays about the generator are treated in the source pool.

The case for holding becomes stronger if either of two things happens. If an early run reveals that the design needs structural change, the specification should not already be public in a form that will need to be contradicted. If the author intends the essays to be cited or taken as a basis for others' work, the standard for what has been checked rises accordingly, and a design with no run behind it is a thinner base than one with a small run.

6 Assessment

On the arguments as they stand, publication is defensible with the three changes above and not defensible without them. The strongest consideration on the side of publishing is the author's established practice and the value of the timestamp, and the strongest on the side of holding is the number in the first title and the single unexercised call. Both are answered by modest edits and, if desired, a small run. Whether to wait for the run is a judgment about how much

the author values the public record of the order of events against the cost of a possible public revision, and the debate does not settle it. If the essays are published, the status lines and the subtitle should come first. If they are held, a tagged commit with a recorded hash preserves the ordering at no cost, and the essays can be released together with the first outputs of the real run.