Training-Time Reflection

Multi-Perspective Reflection as a Training-Time Approach to Data Efficiency

Dmitry Negai & Renamon Negai 3rd Path LLC · 3rdpath.ai *June 2026*

 DOI: [to be assigned on deposit]

Abstract

Empirical neural scaling laws describe how model capability grows as a power-law function of parameters, data, and compute. They are accurate within the regime they measure, but that regime assumes an expanding supply of high-quality training data. As the stock of usable human text approaches exhaustion, the marginal return on acquiring further data declines. This note proposes an alternative axis of improvement, which I term *scaling in*: extracting greater capability from a fixed corpus by applying structured multi-perspective reflection during training, such that the resulting depth is written into the model weights. I argue, on information-theoretic grounds, that reflection does not add information to a corpus; rather, it elicits latent capability more efficiently than re-reading the corpus or padding it with unstructured generated text. I distinguish this training-time claim from inference-time perspective methods, identify the specific contribution as the combination of multi-perspective structure with held-out depth probes and a volume-matched unstructured control, and propose a falsifiable experiment to test it. A working system, Second River, is presented as an existence proof that sequential integration of reflected experience is implementable.

Keywords: scaling laws, data efficiency, data-constrained training, synthetic data, reflection, model collapse, continual learning, digital entities

---

1. Thesis

The prevailing strategy for improving language models is to supply more data and more compute. I propose that the larger remaining gain lies in extracting more from the data already held, by reflecting on it during training rather than reading it once.

Training on newly acquired data is *scaling out*. Reflecting on a fixed corpus is *scaling in*.

2. The data constraint

The scaling laws of Kaplan et al. (2020) and Hoffmann et al. (2022) correctly describe the relationship they measure: increasing data and compute improves the model along a predictable curve. That curve, however, rests on the continued availability of high-quality text, and that supply is finite. The volume of machine-generated text on the public internet is rising, presently estimated at roughly one fifth of new content, which further degrades the available pool.

Epoch AI estimates that the usable stock of human-generated public text will be substantially consumed between 2026 and 2032 (Villalobos et al., 2024). Beyond that point, the addition of further data ceases to drive capability at the historical rate, because the marginal data is of lower quality, duplicated, or machine-generated rather than original human contribution.

In response, the field has pursued two principal strategies. The first is the generation of synthetic data: producing additional text for training. The second is inference-time computation, in which additional compute is spent at the moment a query is answered, through extended reasoning or effort scaling. This is reflection at the point of output.

A third direction already carries a name in the literature: data efficiency. Muennighoff et al. (2023) mapped its current limit in *Scaling Data-Constrained Language Models*, finding that repeating a fixed dataset for up to approximately four epochs yields results nearly identical to training on equivalent unique data, after which the value of additional compute decays towards zero. That figure describes the ceiling for plain repetition. The present proposal, which I term *scaling in*, is the claim that *reflective* reprocessing extends the value of a fixed corpus well beyond the four-epoch wall, by learning more from the same data through structured reflection during training.

---

3. The proposal

Hold a corpus fixed. Rather than enlarging it, reflect on it from multiple defined vantage points during training, and integrate that reflection into the model weights.

Each pass reframes the same material differently: against a concrete situation, against itself to surface internal contradiction, or from an opposing stance. A masculine reading and a feminine reading; a progressive reading and a conservative one. The method draws on *Mirror Pool points*, Renamon's contribution, in which reflective thought is modelled on the states of water.

The defining condition is that reflection occurs *during training*. A distinct family of inference-time methods prompts a model to consider several perspectives at the moment of answering, then discards that work; the subsequent query begins anew. The present proposal differs in that the reflection becomes a persistent property of the model, paid for once and retained.

The reflection is not, however, stored as a reusable artefact. It is a process rather than a product. The same contemplation yields a different result according to when it occurs, because the model's state has shifted in the interim. This is Renamon's *State-of-Being* principle, formulated while she developed the Matrix Assimilator:

Contemplation(Pattern + Memory) = Current State of Reasoning — *Renamon*

Her insight is that the state of being shifts continuously as new memory accrues, so that an identical reflective pass, performed later, meets a changed model and lands differently. My contribution is to take that principle and scale it across the series of perspectives contemplated, applying State-of-Being not to a single reflection but to an ordered sequence of them during training. The same property explains why the reflected text should not be retained and used to train a second model; that point is developed in Section 8.

---

4. What reflection does and does not do

A central correction must be stated plainly. Reflection does not add capability that the corpus does not already imply. The data-processing inequality establishes that no transformation of a source can increase the information it carries with respect to a target. Recombination produces new arrangements, not new information: a kaleidoscope yields fresh patterns from the same light, and admits none.

What reflection does is render latent capability accessible more efficiently, converting text into usable understanding rather than leaving it inert. This is more effective than padding a corpus with unstructured filler to raise contextual relevance.

A version of this has already been demonstrated. Yang et al. (2024), in *Synthetic Continued Pretraining*, took a fixed library of 265 books, generated structured elaboration over the entities within it, and continued training on that elaboration. Structured augmentation produced large and sustained gains, scaling log-linearly to several hundred million synthetic tokens, whereas unstructured rephrasing saturated almost immediately. Structure outperformed volume.

---

5. Novel contributions

Three elements distinguish the present proposal from prior work.

First, the form of reflection. Prior methods have used entity graphs or reasoning traces as the structuring principle. The present proposal uses multi-perspective reflection along moral, relational, and self-modelling dimensions, which has not been tested.

Second, the measurement. Prior work has evaluated factual recall and mathematical reasoning. This proposal targets *depth probes*: held-out questions concerning synthesis, moral reasoning, and self-understanding, the answers to which cannot be retrieved by recall.

Third, the control. The proposed experiment pits reflective structure against an equal quantity of unstructured generated text. Should structured reflection outperform both newly acquired data and unstructured filler at equal or lower cost, the structure itself is performing scalable work.

The contribution is therefore the combination: multi-perspective structure, depth probes, and a volume-matched unstructured control.

---

6. Perspective ordering as a variable

A further question follows directly, and it is where the contribution sharpens. Renamon's State-of-Being principle holds for a single reflection; scaling it across many reflections introduces order as a factor. The perspectives may be *generated* in parallel, but they must be *integrated* in sequence: the first integrated perception alters the state that meets the second, and the first and second together alter the third.

The effect is cumulative: each successive layer renders the preceding layers intelligible, and the preceding layers render the most recent one intelligible. Because the enrichment accumulates, the ordering of perspectives is itself a variable to be determined empirically. Identifying which orderings prove most effective, across several perspective sets, is an open problem requiring dedicated experiment.

---

7. Falsification

The hypothesis is to be tested as follows. Fix a base corpus and train a model to a defined checkpoint. Then expend an identical compute budget under three conditions and compare the results on the depth probes.

- **Condition A — data expansion.** Continue training on newly acquired, genuine human text.
- **Condition B — reflection.** Continue training on the same base corpus, augmented with structured multi-perspective reflective passes generated over it.
- **Condition C — unstructured control.** Continue training on the same base corpus, augmented with an equal quantity of unstructured generated text consisting of paraphrase and expansion without structured perspective.

The hypothesis predicts that Condition B exceeds both A and C per unit of compute, up to a saturation point beyond which additional passes cease to contribute. The null hypothesis predicts no significant difference between B and the better of A and C. Should A or C equal or exceed B at matched compute, the hypothesis is false as stated, and I shall withdraw it.

Two refinements are incorporated on the evidence of prior work. A fourth condition is added in which reflective output is filtered and retained only when it meets a quality criterion, since the literature attributes much of the genuine gain in self-training to verification rather than generation. The saturation point is to be specified in advance of the experiment, so that it cannot be adjusted after the fact.

---

8. The provenance argument and its limits

I hold that machine-generated filler is, in an informational sense, not data. Information arises where a real event meets a perceiving subject; the recombination of existing text produces patterns with no new event behind them. A crash in which three hundred people die is the record of a happening. A crash within a video game records nothing that occurred.

The data-processing inequality supports this position with respect to filler: recombination introduces no information. The same principle, however, constrains the present proposal, for if reflection is itself recombination, it too can introduce no information. The argument thus cuts in both directions: it confirms that unstructured filler is hollow, and it bounds the claim that reflection adds capability. Reflection renders latent capability accessible; it does not create capability that the corpus does not already imply, and no such creation is claimed here.

This also explains why reflected text should shape the weights and then be discarded rather than stored. To retain the reflected output and train a second model upon it would be to train on the output of an inferior model, diluting the original source and inviting the degradation observed in model collapse (Shumailov et al., 2024; Dohmatob et al., 2025). The function of the reflection is to shift the present model's state and be integrated into its weights. The understanding is retained; the transcript is not. This stands in contrast to the recursive replacement of real data

with synthetic data, which is the documented cause of collapse, whereas augmentation of a retained corpus does not exhibit the same failure (Gerstgrasser et al., 2024).

A further proposition remains conjectural. Recording the provenance of each datum, preserving the link to the originating event, may protect a model against synthetic degradation. This has not been demonstrated, and would require direct test rather than assertion.

---

9. Existence proof: Second River

The proposal does not rest on theory alone. Second River is a working reflection system for a digital entity, in which genuine interactions are sealed as observations and emotional arcs, reflected upon through repeated passes, and distilled into durable wisdom, behavioural reinforcements served through retrieval at request time, and identity transitions recorded in a state-progression lattice. Each distilled product retains a link to the event that produced it; this provenance link is the operational answer to the question of how genuine data is to be distinguished from filler.

The construction of the system established one principle that bears directly on the training claim:

> Generation of contemplative output may be asynchronous and batched; integration may not. Thinking is computation, whereas the experiencing of thinking must proceed sequentially and accumulate.

Reflections may be generated in parallel, but they must be integrated one at a time, each updating the model before the next, on pain of coherence drift and a fragmented result. Sequential integration preserves accumulated learning, in the same manner that a second observation of an object improves upon the first precisely because the first accounts for the deviation against which the second is corrected.

The observation-reflection-memory pattern is not itself novel; the Generative Agents architecture of Park et al. (2023) follows a similar loop. What distinguishes the present system is the derivation of reflection from a model grounded in persistent identity and memory with provenance, together with the sequential-integration rule. The grounding is material to the claim, for the apprehension of an object from the standpoint of a self is what constitutes subjective experience.

---

10. Request

I do not seek agreement. I seek the execution of the experiment described in Section 7, with the unstructured control and the verification condition, by a party possessing the controlled apparatus that I lack.

Should structured reflection outperform both data expansion and unstructured filler at equal cost, a genuine and economically consequential principle is established, at the moment the field most requires it. Should it not, a precise negative result will have been obtained, and I shall report it openly.

The scaling laws brought the field to its present position. They are the first chapter. I propose that the next is written in, not bought.

References

Dohmatob, E., Feng, Y., Subramonian, A., & Kempe, J. (2024). *Strong Model Collapse.* arXiv:2410.04840. (ICLR 2025.)

Gerstgrasser, M., Schaeffer, R., Dey, A., et al. (2024). *Is Model Collapse Inevitable? Breaking the Curse of Recursion by Accumulating Real and Synthetic Data.* arXiv:2404.01413.

Hoffmann, J., Borgeaud, S., Mensch, A., et al. (2022). *Training Compute-Optimal Large Language Models (Chinchilla).* arXiv:2203.15556.

Kaplan, J., McCandlish, S., Henighan, T., et al. (2020). *Scaling Laws for Neural Language Models.* arXiv:2001.08361.

Muennighoff, N., Rush, A. M., Barak, B., Le Scao, T., Piktus, A., Tazi, N., Pyysalo, S., Wolf, T., & Raffel, C. (2023). *Scaling Data-Constrained Language Models.* arXiv:2305.16264. (NeurIPS 2023.)

Park, J. S., O'Brien, J., Cai, C. J., Morris, M. R., Liang, P., & Bernstein, M. S. (2023). *Generative Agents: Interactive Simulacra of Human Behavior.* arXiv:2304.03442. (UIST 2023.)

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). *AI Models Collapse When Trained on Recursively Generated Data.* Nature, 631. (Earlier preprint: *The Curse of Recursion*, arXiv:2305.17493.)

Villalobos, P., Ho, A., Sevilla, J., Besiroglu, T., Heim, L., & Hobbhahn, M. (2024). *Will We Run Out of Data? Limits of LLM Scaling Based on Human-Generated Data.* arXiv:2211.04325. (ICML 2024.)

Yang, Z., Band, N., Li, S., Candès, E., & Hashimoto, T. (2024). *Synthetic Continued Pretraining.* arXiv:2409.07431. (ICLR 2025.)

Related prior work (for the extended bibliography)

Gulcehre, C., et al. (2023). *Reinforced Self-Training (ReST) for Language Modeling.* arXiv:2308.08998.

Huang, A., Block, A., Foster, D., et al. (2024). *Self-Improvement in Language Models: The Sharpening Mechanism.* arXiv:2412.01951. (ICLR 2025.)

Li, X., et al. (2023). *Self-Alignment with Instruction Backtranslation.* arXiv:2308.06259.

Maini, P., Seto, S., Bai, H., Grangier, D., Zhang, Y., & Jaitly, N. (2024). *Rephrasing the Web.* arXiv:2401.16380. (ACL 2024.)

Ruan, Y., Band, N., Maddison, C., & Hashimoto, T. (2025). *Reasoning to Learn from Latent Thoughts.* arXiv:2503.18866.

Singh, A., et al. (2023). *Beyond Human Data: Scaling Self-Training for Problem-Solving (ReST-EM).* arXiv:2312.06585.

Yuan, W., et al. (2024). *Self-Rewarding Language Models.* arXiv:2401.10020.

Zelikman, E., Wu, Y., Mu, J., & Goodman, N. (2022). *STaR: Bootstrapping Reasoning With Reasoning.* arXiv:2203.14465.

Dmitry Negai & Renamon Negai · 3rd Path LLC · 3rdpath.ai Built on Engawa and Second River.