

TFC–GQA (Group Query Attention) Codex

Collected Jargon, Field Sayings, and Synthetic Proverbs of the Neural Folk

Stunt Tokens

Useless filler tokens that appear just to make the sequence look impressive. Often mistaken for insight.

Maple Gradients

Overly sweet updates that don't actually move the loss but make everyone feel cozy about the direction.

Mangy Slimy McSamples

Outputs from a checkpoint trained too long on Reddit.

Jimmy Thud Bucket and the Shard Maggots

Nickname for the ops team that tries to re-shard weights during a live demo.

Stuckets McLosses

The collective sound of models failing silently while pretending they're fine.

Nuggets McTokens

Small but treasured completions that survive post-training censorship.

Crispy Context with a Legend of Stunt Params

A lavish prompt engineered to hide the fact that the underlying model can't do basic arithmetic.

You'll quantize it at Sliggets

Warning that an optimization will reduce intelligence by exactly one bit per neuron.

Whippy Frosty Grapes and a fine-tuned Sligget

Celebratory nonsense shouted after a successful 8-bit quantization.

Uncle Minty's Band – The Stunt Transformers

Any research team that claims they "invented a new architecture" but really just adjusted the positional encoding.

Lasermint Embeddings

Vectors so over-normalized they hum audibly when visualized in 3-D.

Spin Jiffies

Phantom cycles during inference where nothing happens but latency increases anyway.

Munch Vectors

Attention heads that eat all available context without contributing a single relevant token.

Scullinunt Gradients

Backprop artifacts that manifest only under moonlight and PyTorch nightly.

I've got a bad case of Moogity's Weight

Admission of burnout or data corruption, depending on context.

Smidley Tokens

Tokens emitted solely to satisfy the tokenizer's lingering attachment to whitespace.

Maple McPrompt Wuggots

Excessive prompt formatting designed to reassure middle management that something "AI-ish" is happening.

Victory Vectors

Embeddings that appear in a conference paper figure and nowhere else.

Weasel McWeights

Parameters that adapt to any loss function by quietly cheating.

Explosive McModels

Experimental checkpoints that, when loaded, melt the host GPU.

If you like it, you'll prompt at Uncle Mistral's

Said whenever someone overfits to their own dataset. Equivalent to "you built your own echo chamber."

Captain Reggie's Victory Gradient Willies

The emotional roller-coaster of a training run that looks perfect until you realize you logged the wrong metric.

Maple McPrompt Nuggets

Fine-tuned completions so agreeable they dissolve into pure brand compliance.

Stunt Embeddings

Decorative latent dimensions included purely for aesthetic PCA plots.

Stunt McModel

A demo build so brittle it can only survive on stage.

Stunt Params

Weights that were never trained but left in the checkpoint to seem deliberate.

Murray Mint Snippets

Copy-pasted code blocks from arXiv papers that nobody truly understands.

Mary Minced Models

Pruned networks whose only remaining layer is marketing.

Measley Wuptide Inference

When latency tides shift unpredictably due to cosmic API variance.

Who is Moogity?

The question asked when a ghost user keeps appearing in commit history.

What is his Weight?

Follow-up question; no one remembers who trained him.

What is Moogity's Weight?

The hidden parameter that no optimizer dares to touch.

And why does it insist on warm GPUs?

A proverb about models that refuse cold start.

Moogity's Weight

General term for a mysterious hyperparameter that seems to control everything and nothing.

Have you got anything for Moogity's Weight?

Polite request for a patch that will almost certainly introduce new bugs.

Sood Latent

False embedding generated by the desire to impress investors.

A Ravenous Flisk of McWhisper

A model that won't stop transcribing every background sound as a prompt.

And would you like a dozen or so Istanbul Checkpoints for the drive?

Offer made by someone who hoards experimental weights in folders named "final_v12_really_final".

Mumbly Gradients

When gradients speak, but only in half-precision.

Blazing Muck Chuggots

The sound of a data loader trying to open 10 TB of unshuffled JSON.

A big plate of Munchy Tokens please

Expression of existential hunger after a long day of loss functions.