DAY 15 · Part 2 · How We Got Here

Why did just making it bigger give it new skills? — 2020: GPT-3

Five-Minute AI Lessons · Sep 21, 2026

Same design, just more size and more reading, and it learned to copy from a few examples.

In a nutshell

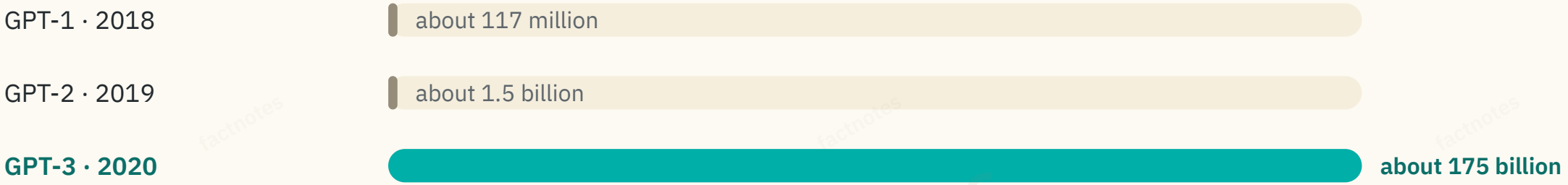
1. They didn't change yesterday's Transformer design at all — they just made it bigger, and in 2020 out came a model with **175 billion** numbers. Its name is GPT-3.

2. The surprise wasn't the size, it was **how it learned**. With no retraining, you could tuck just two or three examples into your question and it would pick up the task right away.

3. That's why the paper was titled **'Language Models are Few-Shot Learners'** — 'few-shot' means from just a few examples. From then on, 'what you tell it,' the prompt, started to really matter.

In one picture

How much GPT grew in 3 years



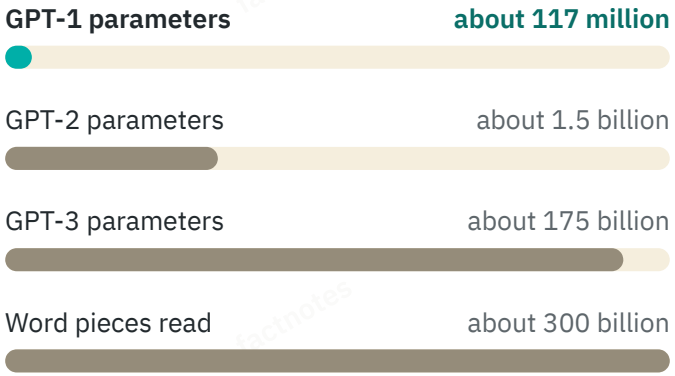
With the same ruler, the first two bars are barely even visible. That's how big the jump was.

Key idea: keep the design, grow only the size and the reading — and new skills showed up

A feel for digits — how big did GPT get?

about 117 million

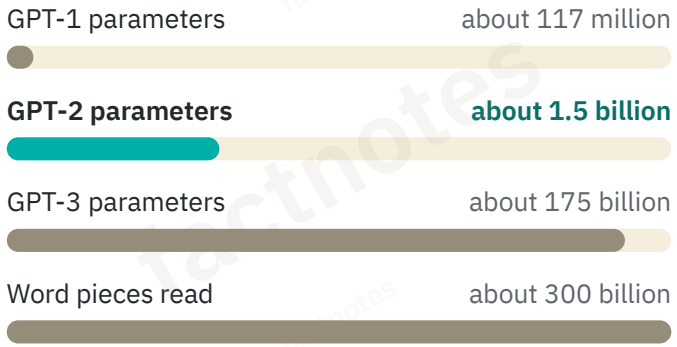
GPT-1 parameters



Back then, this much counted as a 'big model.'

about 1.5 billion

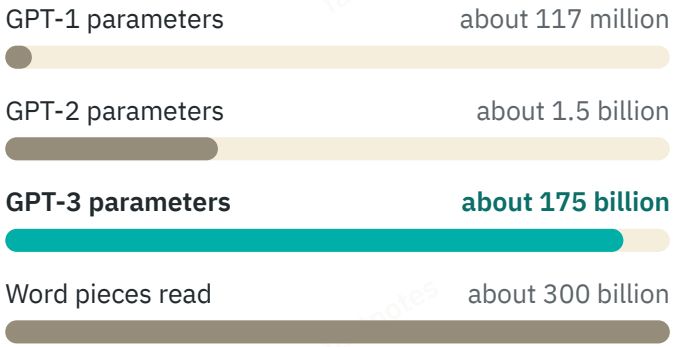
GPT-2 parameters



In just one year, it jumped by more than one digit.

about 175 billion

GPT-3 parameters



That's more than 100 times GPT-2. Same design — they just made it this much bigger.

Move the slider to guess 'about 10 to the what?' first. Then reveal the answer to see the real number. The goal...

Try it yourself in the web lesson · next-token.net/en/learn/day-15/

Quick check

Just three questions

Q1

What made GPT-3 most different from the models before it?

① It searched the internet in real time
② Show it a few examples, and it did a new task
③ People put in the rules one by one
④ It switched its own power on and off

Q2

About how many numbers (parameters) does GPT-3 have?

① About 170 million
② About 1.7 billion
③ About 175 billion
④ About 17.5 billion

Q3

What do you call what you give the model, examples tucked in and all?

① An update
② A download
③ A refresh
④ A prompt

Know the why, too

Q1

② With no retraining, it just looked at the examples in the question and followed them. That's 'few-shot' learning — which is where the paper's title came from.

Q2

③ It's 175 billion. GPT-2 was about 1.5 billion, so it grew more than 100 times.

Q3

④ Everything you hand to a model is called the prompt. Starting with GPT-3, how you wrote it could change the results a lot.

Same design, just more size and more reading, and it learned to copy from a few examples.



Tomorrow: 2022: ChatGPT — how a chat window, more than the model itself, changed the world.