OUTPUT PER FRONTIER AI MODEL (2026)

	2021 (GPT-3)	2024 (GPT-4o)	2025 (Frontier LLM)	*Human output* 1×	*10×*	*100×*
Tokens per day	6B	200B	12T	176T	1.7Qa	17Qa
Words per minute	3.1M	104M	6.2B	92B	922B	9.2T
Tokens per second	70K	2.3M	139M	2B	20B	204B
New York Public Libraries per day	0.004	0.125	7.5	110	1,100	11,000
1 New York Public Library per *x*...	266 days	8 days	3.2 hours	13.1 mins	1.3 min	7.9 secs

- 25/Mar/2021: 'We currently generate an average of 4.5 billion words per day' https://openai.com/blog/gpt-3-apps/
- 23/Jul/2024: 'GPT-4o mini... processing more than 200B tokens per day' https://x.com/sama/status/1815437745550172617
- 19/Dec/2025: 'Let's say that an AI company today might be generating something on the order of 10 trillion tokens a day. It's more, but it's not a quadrillion tokens for anybody, I don't think.' https://youtu.be/2P27Ef-LLuQ?t=1861
- 19/Dec/2025: 'We're going to have these models at a company outputting more tokens per day than all of humanity put together. And then 10 times that, and then 100 times that...' https://youtu.be/2P27Ef-LLuQ?t=1909

Figures rounded for readability. 1 word ≈ 1.33 tokens. 133K tokens ≈ 100K words ≈ 1 book. 1 New York Public Library ≈ 12M books ≈ 1.6T tokens ≈ 1.2T words. Human output ≈ 16K words x 8.3B people ≈ 132.8T words ≈ 176.6T tokens. Alan D. Thompson. February 2026.