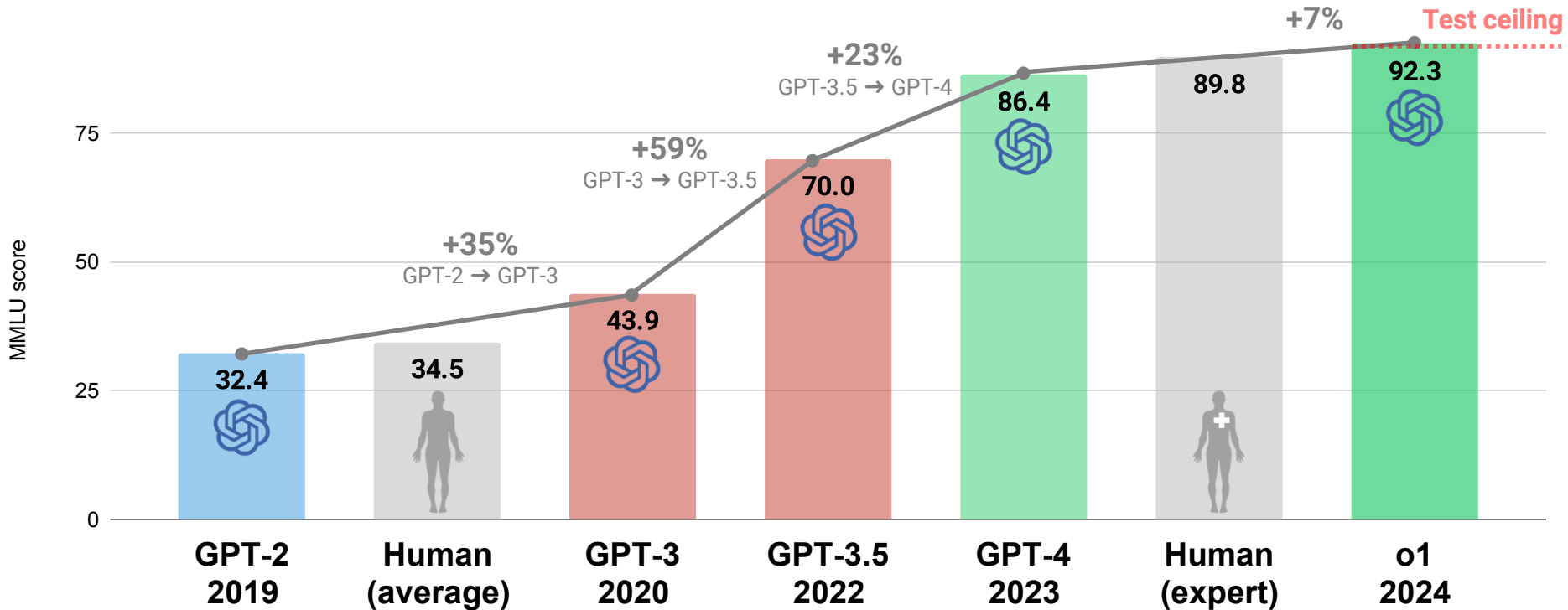


LLMs: SMARTER THAN WE THINK (SEP/2024)



MMLU (Massive Multitask Language Understanding) benchmark features 57 tasks including mathematics, US history, computer science, law, and more. % increases rounded. <https://life architect.ai/o1/> Alan D. Thompson. 2024.

