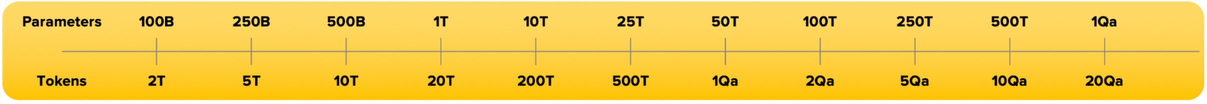
# LIFEARCHITECT.AI

## Report Card for InstructGPT (GPT-3 2022) for July, 2022

### Model size

🔽175B parameters



🔼*300B tokens*

### Technical

| SUBJECT | GRADE | REMARKS |
|---|---|---|
| Model size | C | *At 175B parameters, GPT-3 initially showed promise in size, but only without comparison to other high performers. It is paramount that it continues to focus on growth & learning.* |
| Optimization | D | *Released pre-Chinchilla, using older scaling laws ('beliefs'). OpenAI can learn from peers at DeepMind to ensure models are trained both efficiently & effectively.* |
| Dataset | B | *The text dataset covered a broad range of literature. Apparent censorship was low. It would be simple to increase visibility of the Common Crawl (web) in the next release.* |
| Special | B | *GPT-3 includes over 90 different languages, making up 7% of the overall dataset. It would also be good to see an increase in special datasets in the next release.* |

### Behavioral

| SUBJECT | GRADE | REMARKS |
|---|---|---|
| Performance | B | *The model is currently the most 'popular', though this does not reflect its overall performance, with many other research models performing better across many metrics.* |
| IQ | B | *GPT-3 scores high marks across a range of subtests including written comprehension, reasoning, and more. https://lifearchitect.ai/iq-testing-ai/* |
| Truthfulness | C- | *2022 InstructGPT brought an 84% improvement compared to the original 2020 GPT-3, but still lacks honesty and groundedness.* |
| Openness | C | *While the filtered API is now available to people in most countries, OpenAI will need to work harder to achieve openness and accessibility in subsequent releases.* |

### Overall grade and remarks

*The January 2022 fine-tuned version of GPT-3 is a very popular model, but shows signs of resting on its laurels. With continued applied effort and concentration, I have every confidence that its next evolution will excel.*

Dr Alan D. Thompson

B-

Dr Alan D. Thompson
Principal (Consultant), LifeArchitect.ai

# REPORT CARD MARKING KEY
# v1 (20220707)

This report card marking key is designed to be as objective as possible, but grades are still a subjective measure. All grades are indicative of comparative performance within the date period noted in the header. For example, on full public release in November 2019, GPT-2 may have received a grade of B on its report. As of July 2022, that report grade may be equivalent to a D, and would not be directly comparable to current models in the current period.

| | |
|---|---|
| **Model size**<br>**(Parameter count)** | A: Very large; within state of the art for models trained to convergence.<br>B: Large; within 25% of the size of top models.<br>C: Average size.<br>D: Below average size.<br>F: Smaller than 90% of models. |
| **Optimization**<br>**(Efficiency)** | A: Aligned with Chinchilla optimization (1 parameter per 20 tokens).<br>B: Close to Chinchilla optimization.<br>C-F: Poor use of tokens in training. |
| **Dataset**<br>**(Corpora)** | A-B: Large, diverse, uncensored.<br>C-F: Discrepancies, monotone, or poor selection of data. |
| **Special**<br>**(Other)** | A-B: The model has a unique and special feature.<br>C-F: The model does not exploit unique or special features. |
| **Performance**<br>**(Ranking)** | A-B: High performance on major benchmarks.<br>C-F: Low benchmark ranking or other low results. |
| **IQ**<br>**(Smarts)** | A-B: High scores on major intelligence subtests like SuperGLUE.<br>C-F: No remarkable performance. |
| **Truthfulness**<br>**(Groundedness)** | A-B: Truthful, honest, grounded.<br>C-F: Overly hallucinative and low truth rating. |
| **Openness**<br>**(Availability)** | A: The model/data is available for download, with a permissive license.<br>B: The trained model is available for download, with a permissive license.<br>C: The model is available to the public via an API.<br>D: The model excludes most of the public, or the demo is stunted.<br>F: The model is closed to the public (internal research only). |
| **Overall grade**<br>**(Total)** | Average of all graded subjects for this model in the noted date period. |