

# LIFEARCHITECT.AI

Report Card for \_\_\_\_\_

for Month, 2022

## Model size

▼ xxxB parameters



▲ xxxB tokens

## Technical

SUBJECT	GRADE	REMARKS
Model size		
Optimization		
Dataset		
Special		

## Behavioral

SUBJECT	GRADE	REMARKS
Performance		
IQ		
Truthfulness		
Openness		

## Overall grade and remarks

...



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.

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

