

Seal TikZ

@nullvaluetensor on twitter

November 2024

1 Prompt used

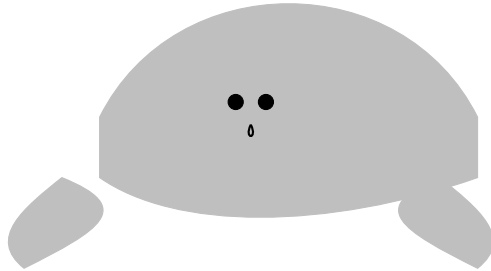
"Draw me a tikz seal."

2 OpenAI models

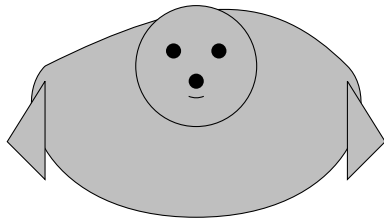
2.1 4o-latest-11-3-2024



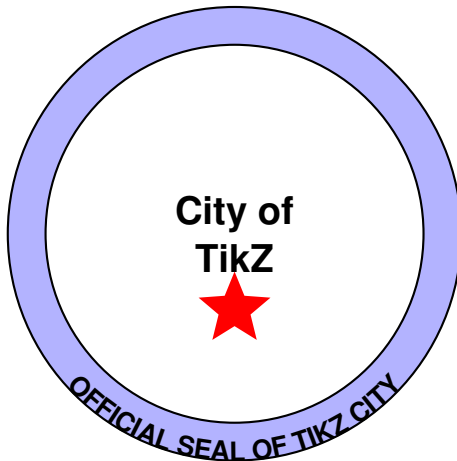
2.2 4o-08-06



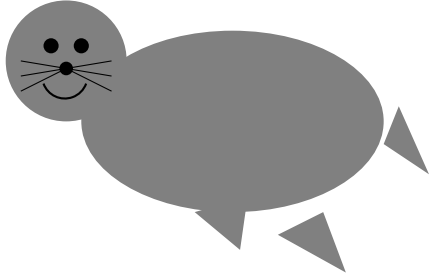
2.3 gpt-4o-0513



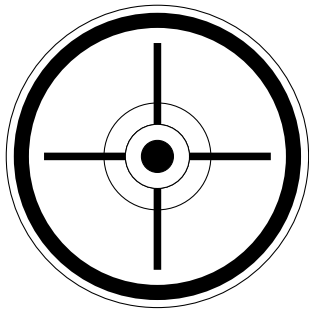
2.4 im-a-good-gpt2-chatbot



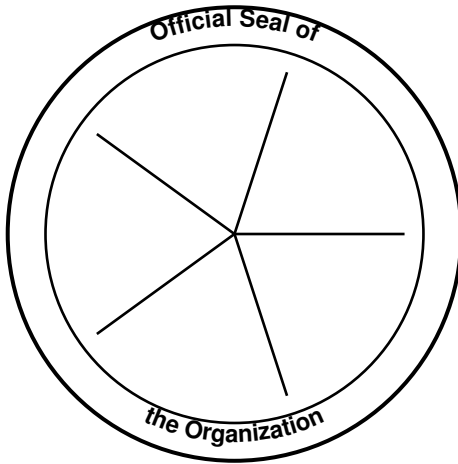
2.5 i-am-also-a-good-gpt2-chatbot



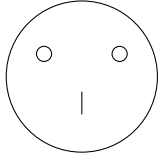
2.6 gpt-3.5-0314



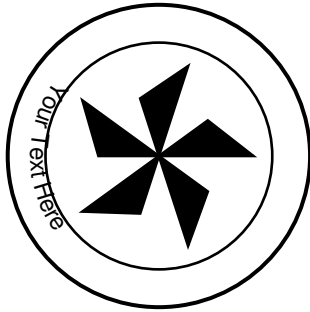
2.7 gpt-4-0314



2.8 gpt-4-0613



2.9 gpt-4-0314-32k



2.10 gpt-4-0613-32k

FAILED

2.11 gpt-3.5-0613

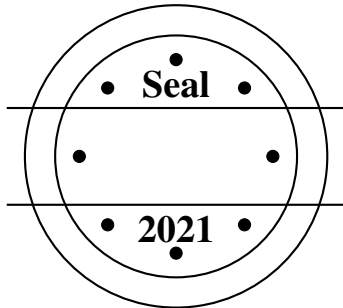
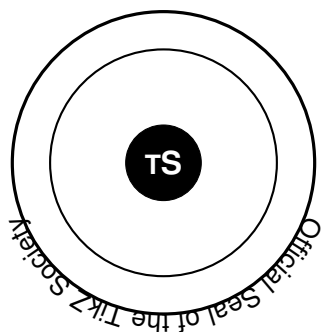


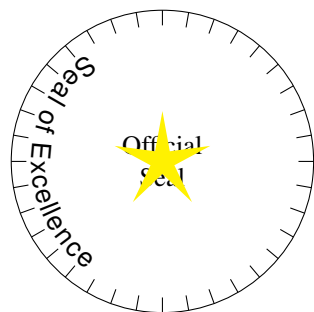


Figure 1: Enter Caption

2.12 gpt-4-turbo-1106-vision



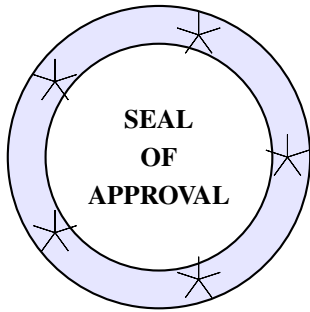
2.13 gpt-4-turbo-0125



2.14 dalle 3

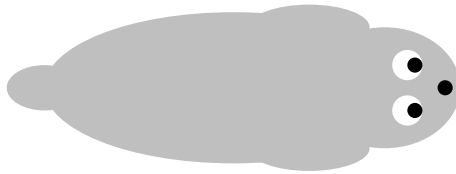
3 XAI models

3.1 grok-2-08-13

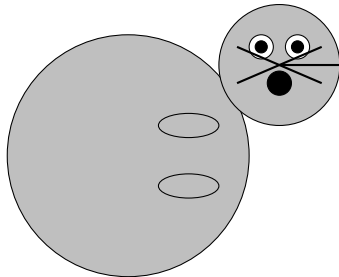


3.2 Anthropic models

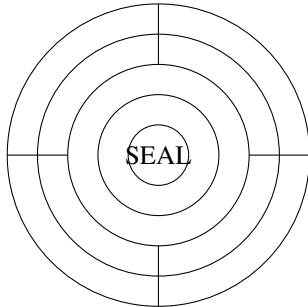
3.3 Claude-3-opus-20240229



3.4 claude-3-sonnet-20240229



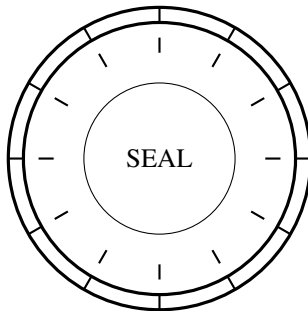
3.5 claude-3-haiku-20240229



3.6 claude-3-5-sonnet-20240620

FAILED (REFUSAL)

3.7 claude-3-5-sonnet-20241022



4 LLaMA variants

4.1 llama-3.1-405b-bf16

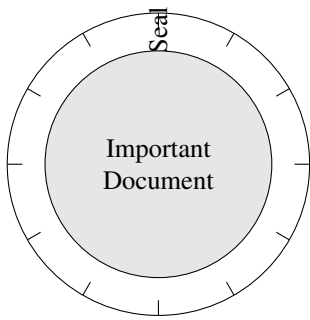


4.2 Hermes-3-405b

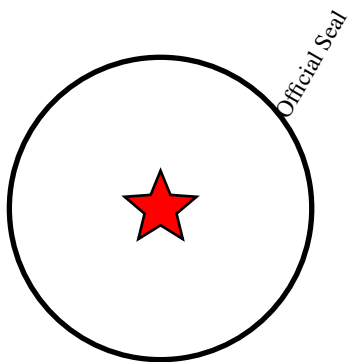


5 Google models

5.1 gemini-1.5-pro-001



5.2 gemini-1.5-pro-002



5.3 gemini-pro-001-1.0

