

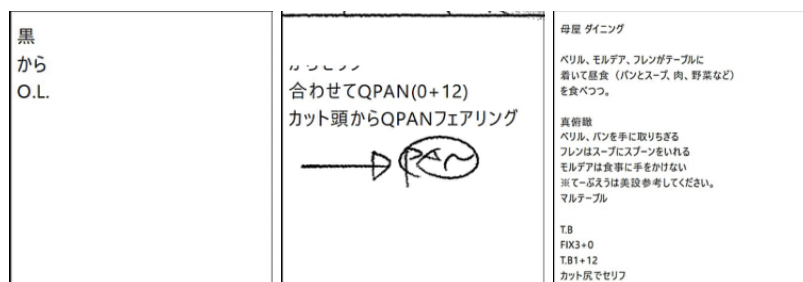
報告書

1 今週の進捗

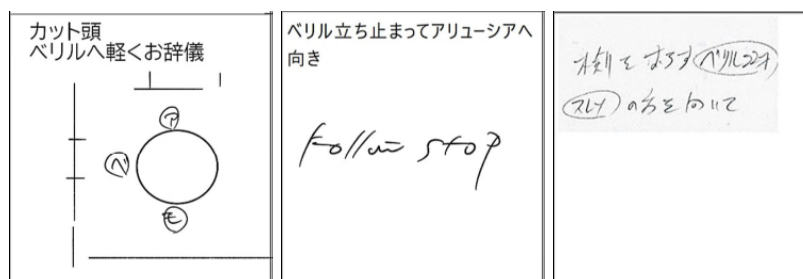
- GPT API の料金調査

2 GPT への入力

適当に選んだ 10 ページで OCR を試した.



(a) page002 ("黒"を含む) (b) page007 (手書きの記号) (c) page018 (小さめの文字)



(d) page024 (形式ずれ) (e) page053 (手書き文字) (f) page060 (手書きページト)



(g) page061 (手書き文字と記号) (h) page082 (イラスト) (i) page094 (文章量少なめ) (j) page105 (量多め)

2.1 GPT API の料金調査

model	料金 (input)	料金 (output)
gpt-4o-mini	0.15	0.60
gpt-4.1-mini	0.40	1.60
gpt-4.1-nano	0.10	0.40

model	料金 (表示)	料金 (計算)
gpt-4o-mini	0.13	0.130472
gpt-4.1-mini	0.01	0.013099
gpt-4.1-nano	0.01	0.003822

gpt-4.1-mini のほうが安い, これは画像入力料金の違いだと考えられる. gpt-4.1-mini より gpt-4o-mini のほうが画像入力料金が安いと考えられる.

2.2 OCR 結果

比較項目	gpt-4o-mini	gpt-4.1-mini	gpt-4.1-nano
page002_row1	黒@から@O.L.	黒@から@O.L.	黒@から@O.L.
page007_row2	ハンセン@合わせて QPAN(0+12)@ カット頭から QPAN フェアリング@ →@PAC	はとつ合わせて QPAN(0+12) カット頭から QPAN フェアリング	合せて QPAN(0+12)@ カット頭から QPAN フェアリング
page018_row1	... カット尻でセリフ	... カット尻	... カット尻でセ
page024_row2	カット頭@ ベリルへ軽くお辞儀	カット頭@ ベリルへ軽くお辞儀	カット頭@ ベルリム軽くお辞儀
page053_row5	ベリル立ち止まって アリュージアへ向き@ follow stop	ベリル立ち止まって アリュージアへ向き@ Follow stop	ベル立ち止まって アリュージアへ向き@ follow stop
page060_row3	本当にベルソナ@ スライの方を向いて	杖をもつ ベルコニカ @ ストレの方を向いて	木をおろすベルコウクリ@ スタのを書いて
page061_row1	6@歳スレナヨリ@ T.U@ O.L(2+0)@ 理味より情感.@ 左手にわかる.	ナビカ PANUP!@ ピ送り (土) → (木) へ@ じっと木鉢を見る@スレ 64	フィット PANUP@ @ ピース!! (上) → (下) → じっと 木を見つめる (セリフ)
page082_row1			"Sorry, I can't assist with that."
page094_row1	ベリル ナメ@ 奥にヘンブリッツ@ 残像エコー風@※欠番	ベリル ナメ@ 奥にヘンブリッツ@ 残像エコー風@※欠番	ベルリナメ@ 奥にヘンブリッツ@ 残像エコー風@※欠番
page105_row4	フラッシュ回想@修練場@ 作画流用+ベリル足し@ 振り切るヘンブリッツに 合わせて@ベリル移動@ ヘンブリッツは スローでベリルは@ ノーマル速度	フラッシュ回想@修練場@ 作画流用+ベリル足し@ 振り切るヘンブリッツに 合わせて@ベリル移動@ ヘンブリッツは スローでベリルは@ ノーマル速度	フラッシュ回想@修練場@ 作画流用+ベル足し@ 振り切るヘンブリッツに 合わせてベリル移動@ ヘンブリッツは スローでベリルは ノーマル速度

3 今後の課題

- "gpt-4.1" で間違えたものを試す
- "gemma-3" モデルのプログラムコード経由での実行における動作の安定化および精度改善
- Mixture of Experts (MoE) アーキテクチャを採用した OCR モデルに関する調査

参考文献