

OpenAI公式から学ぶ！ o1モデルのプロンプトエンジニアリング

KINTOテクノロジーズ株式会社

2024/09/18

ついに出た！OpenAIの最新モデル「o1」って何がすごいの？みんなで勉強しよう



プロフィール



KINTOテクノロジーズ株式会社
IT/IS部 生成AI活用PJT 生成AIエバンジェリスト

和田 颯馬 ([@cognac_n](#))

2022年2月入社。生成AI活用PJT所属。内製アプリケーションのレコメンドアルゴリズム作成や、社内外の分析レポート作成などに従事。直近は、社内のLLM活用を推進する業務に注力中。

Index

目次

- 1 KINTOテクノロジーズとは
- 2 o1の立ち位置って？（私見含む）
- 3 o1モデルのプロンプトエンジニアリング
- 4 まとめ

1

KINTOテクノロジーズとは



トヨタ自動車株式会社

トヨタファイナンスサービス株式会社

海外販売金融
現地事業会社

世界40以上の国と地域で
サービスを展開

トヨタファイナンス
株式会社

販売金融・クレジット
カードなど

KINTO
Technologies

KINTOテクノロジーズ
株式会社

KINTO

株式会社KINTO

クルマにかかる諸費用がコミコミ！

他に必要なのは「ガソリン代・駐車場代・高速代」など



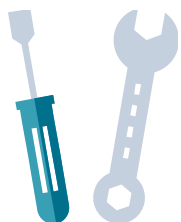
自動車税



自動車保険



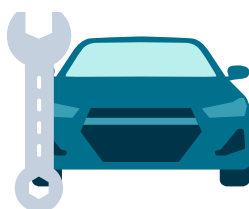
車検



メンテナンス



消耗品



故障修理・代車



車両代 + 登録諸費用

KINTOテクノロジーズが関わっているプロダクト



KINTO
ONE



KINTO
FACTORY

KINTO ONE 中古車



Prism_{Japan}



TOYOTA
wallet

TOYOTA
WOVEN CITY

GLOBAL KINTO
ID Platform

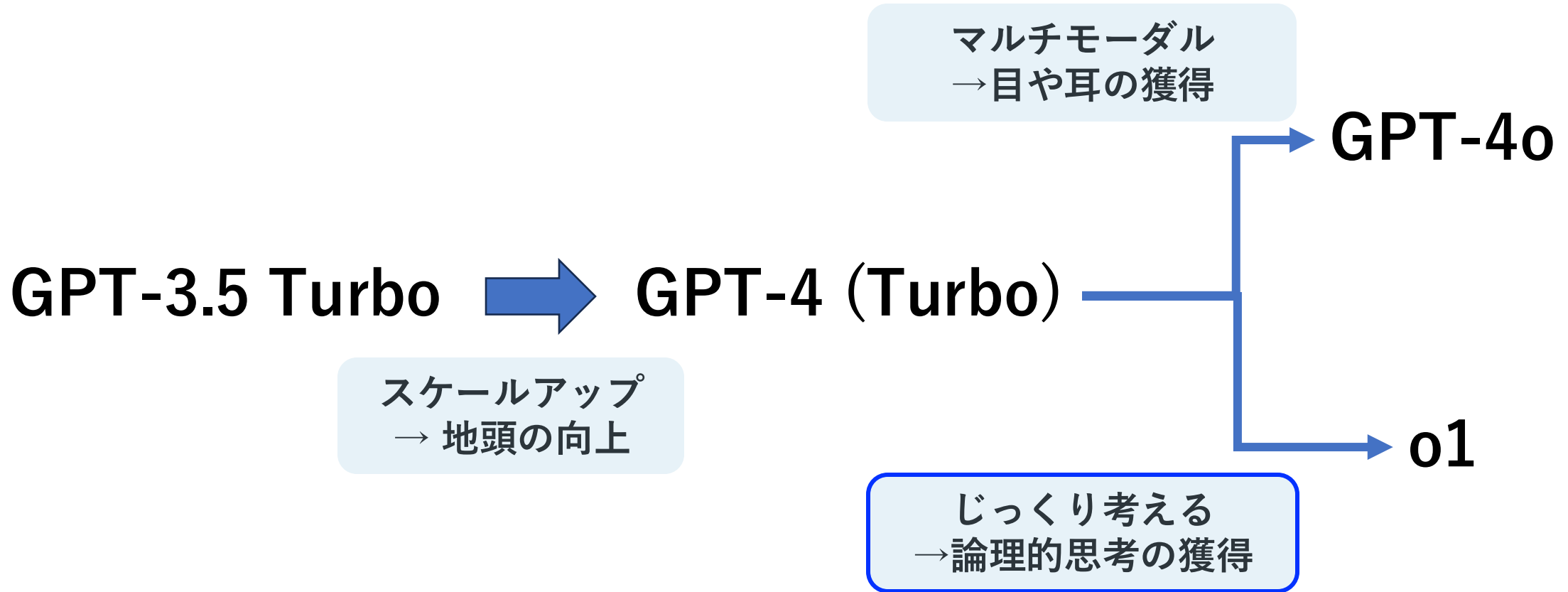
route
by KINTO

モビリティ
マーケット

2

o1の立ち位置って？（私見含む）





※モデルユーザー目線での私見/印象です

GPT-4 (Turbo)

GPT-4o

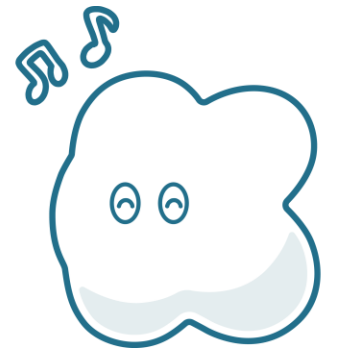
o1

スケールアップ
→ 地頭の向上

マルチモーダル
→ 目や耳の獲得

じっくり考える
→ 論理的思考の獲得

???



※モデルユーザー目線での私見/印象です

GPT-4 (Turbo)

GPT-4o

o1

スケールアップ
→ 地頭の向上

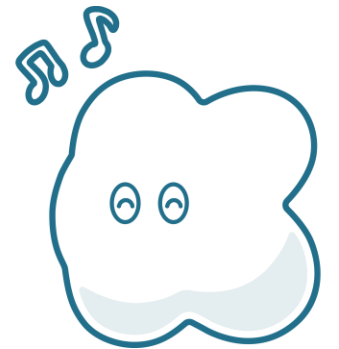
いつか出てくるとてつもない技術に
備えさせようとしている??

マルチモーダル
→ 目や耳の獲得

???

じっくり考える
→ 論理的思考の獲得

今のうちから
使い方に慣れておく



※モデルユーザー目線での私見/印象です

3

o1モデルのプロンプトエンジニアリング



参照した記事

Introducing OpenAI o1-preview

<https://openai.com/index/introducing-openai-o1-preview/>

OpenAI o1 System Card

<https://openai.com/index/openai-o1-system-card/>

Learning to Reason with LLMs

<https://openai.com/index/learning-to-reason-with-llms/>

Using reasoning for routine generation

https://cookbook.openai.com/examples/o1/using_reasoning_for_routine_generation

Using reasoning for data validation

https://cookbook.openai.com/examples/o1/using_reasoning_for_data_validation

Reasoning models

<https://platform.openai.com/docs/guides/reasoning>

使うモデル

- 1 **o1-preview**: 世界についての幅広い一般知識を使用して難しい問題について推論するように設計された、O1モデルの初期プレビュー。
- 2 **o1-mini**: o1のより高速で安価なバージョン。広範な一般知識を必要としないコーディング、数学、科学のタスクに特に適しています。

o1モデルは推論において大きな**進歩**をもたらしますが、**すべてのユースケースでGPT-4oを置き換えることを意図したものではありません。**

画像入力、関数呼び出し、または一貫して高速な応答時間を必要とするアプリケーションの場合、GPT-4oおよびGPT-4o miniモデルは引き続き適切な選択です。ただし、深い推論を必要とし、より長い応答時間に対応できるアプリケーションの開発を目指している場合は、o1モデルが最適な選択肢になる可能性があります。これらのモデルを使用してどのようなものを作成するか楽しみにしています。

<https://platform.openai.com/docs/guides/reasoning>

プロンプトのコツ

Advice on prompting

これらのモデルは、簡単なプロンプトで最もよく機能します。少数のプロンプトやモデルに「ステップごとに考える」ように指示するなどのプロンプト エンジニアリング手法では、パフォーマンスが向上しない可能性があり、パフォーマンスが妨げられることもあります。次に、ベスト プラクティスをいくつか示します。

- **プロンプトはシンプルかつ直接的なものにしてください。**モデルは、詳細なガイダンスを必要とせずに、簡潔で明確な指示を理解して応答することに優れています。
- **思考の連鎖を促すプロンプトは避けてください。**これらのモデルは内部で推論を実行するため、「段階的に考える」または「推論を説明する」ように促す必要はありません。
- **明確にするために区切り文字を使用する:**三重引用符、XML タグ、セクション タイトルなどの区切り文字を使用して、入力の異なる部分を明確に示し、モデルがさまざまなセクションを適切に解釈できるようにします。
- **検索拡張生成 (RAG) での追加コンテキストを制限する:**追加のコンテキストまたはドキュメントを提供する場合は、モデルが応答を過度に複雑にしないように、最も関連性の高い情報のみを含めます。

余計なことをするな

必要な情報だけ
わかりやすく提供せい



素人は黙っとれ…？



<https://platform.openai.com/docs/guides/reasoning>

公式のサンプルを見てみよう！

```
prompt = ""
Instructions:
- Given the React component below, change it so that nonfiction books have red
  text.
- Return only the code in your reply
- Do not include any additional formatting, such as markdown code blocks
- For formatting, use four space tabs, and do not allow any lines of code to
  exceed 80 columns

const books = [
  { title: 'Dune', category: 'fiction', id: 1 },
  { title: 'Frankenstein', category: 'fiction', id: 2 },
  { title: 'Moneyball', category: 'nonfiction', id: 3 },
];

export default function BookList() {
  const listItems = books.map(book =>
    <li>
      {book.title}
    </li>
  );

  return (
    <ul>{listItems}</ul>
  );
}
```

- ・以下のReactコンポーネントを、ノンフィクションの本のテキストが赤くなるように変更してください。
- ・返信にはコードのみを含めてください。
- ・マークダウンのコードブロックのような追加のフォーマットは含めないでください。
- ・フォーマットには4スペースのタブを使用し、コードの1行が80文字を超えないようにしてください。

<https://platform.openai.com/docs/guides/reasoning>

公式のサンプルを見てみよう！

```
prompt = """
I want to build a Python app that takes user questions and looks them up in a
database where they are mapped to answers. If there ia close match, it retrieves
the matched answer. If there isn't, it asks the user to provide an answer and
stores the question/answer pair in the database. Make a plan for the directory
structure you'll need, then return each file in full. Only supply your reasoning
at the beginning and end, not throughout the code.
"""
```

私は、ユーザーからの質問を受け取り、それをデータベースで検索するPythonアプリを作りたいと考えています。そのデータベースでは、質問が答えと対応付けられています。もし近い一致があれば、その一致した答えを取得します。もし一致しなければ、ユーザーに答えを提供してもらい、その質問/答えのペアをデータベースに保存します。ディレクトリ構造を計画し、各ファイルの内容を完全に書いてください。コードの途中ではなく、最初と最後に理由を説明してください。

<https://platform.openai.com/docs/guides/reasoning>

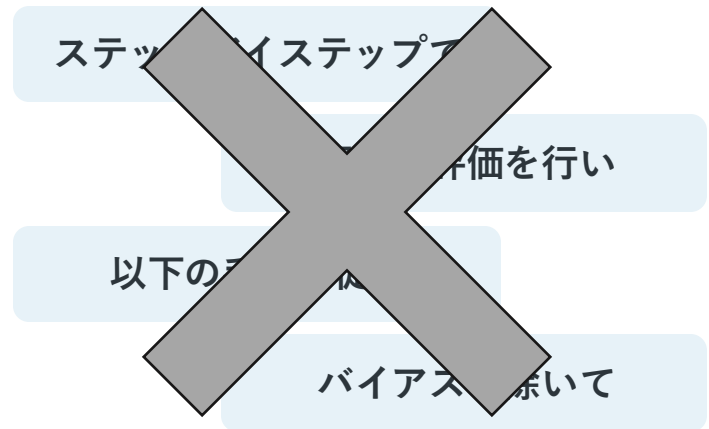
公式のサンプルを見てみよう！

```
prompt = """  
What are three compounds we should consider investigating to advance research  
into new antibiotics? Why should we consider them?  
"""
```

新しい抗生物質の研究を進めるために調査すべき化合物を3つ挙げてください。
それらを考慮すべき理由は何ですか？

→ 何をしたいか、端的に伝える
余計なことはしない

<https://platform.openai.com/docs/guides/reasoning>



GPT-4oとo1-previewで動作を比較

新しい抗生物質の研究を進めるために調査すべき化合物を3つ挙げてください。
それらを考慮すべき理由は何ですか？

新しい抗生物質の研究を進めるために検討すべき3つの化合物は以下の通りです。

1. テイクトニン (Teixobactin)

理由: テイクトニンは、グラム陽性菌に対して非常に有効で、既存の抗生物質に耐性を持つ細菌に対しても効果が確認されています。特に、メチシリン耐性黄色ブドウ球菌 (MRSA) や結核菌に対して有望であり、耐性菌問題の解決に貢献できる可能性があります。

2. ダプトマイシン (Daptomycin)

理由: ダプトマイシンは、膜を破壊することで殺菌効果を発揮するユニークな作用機序を持っています。このメカニズムにより、他の抗生物質に耐性を持つ細菌に対しても効果が期待され、特に複雑な皮膚感染症や菌血症の治療に重要な役割を果たしています。

3. リシン誘導体 (Lysin Derivatives)

理由: リシンはバクテリオファージが産生する酵素で、細菌の細胞壁を破壊する働きを持っています。抗生物質に対する耐性が問題となっている現在、リシン誘導体は、細菌に直接作用する新しい治療法として注目されています。特に、グラム陰性菌に対する有効性が期待されており、新たな感染症治療の道を開く可能性があります。

これらの化合物は、抗生物質耐性菌の問題に対処するために、革新的なメカニズムを持つことが期待されるため、研究を進める価値が高いです。

GPT-4o

<https://platform.openai.com/docs/guides/reasoning>

思考時間: 10 秒

新しい抗生物質の研究を進めるために検討すべき3つの化合物とその理由は以下の通りです。

- 1. テイコバクチン (Teixobactin)**: テイコバクチンは2015年に発見された新しいクラスの抗生物質で、グラム陽性菌に対して有効です。特に、多剤耐性菌であるメチシリン耐性黄色ブドウ球菌 (MRSA) や結核菌に効果があります。この化合物は細胞壁の合成を阻害するため、細菌が耐性を獲得しにくいとされています。
- 2. 抗菌ペプチド (Antimicrobial Peptides, AMPs)**: 抗菌ペプチドは自然界に広く存在し、生体の免疫システムの一部として機能します。広範な抗菌スペクトルを持ち、細胞膜を直接攻撃するため、細菌が耐性を獲得しにくい特徴があります。また、抗ウイルスや抗真菌効果も持つため、多面的な治療が期待できます。
- 3. クオラムセンシング阻害剤 (Quorum Sensing Inhibitors)**: クオラムセンシングは細菌間の情報伝達機構で、病原性やバイオフィーム形成に関与しています。このシステムを阻害する化合物を研究することで、細菌の増殖を直接抑制せずに病原性を低下させる新しいアプローチが可能です。これにより、耐性菌の出現を抑える潜在的なメリットがあります。

これらの化合物は、異なる作用機序を持ち、抗生物質耐性問題の解決に寄与する可能性があります。そのため、さらなる研究と開発が重要です。

o1-preview

GPT-4oとo1-previewで動作を比較

クルマのサブスクサービスを利用していただいているお客様の満足度を向上させるために重要な施策を3つ考案してください。それらの施策を選んだ理由は何ですか？

🌀 クルマのサブスクサービスの満足度を向上させるために、以下の3つの施策が効果的だと考えます。

1. パーソナライズされた提案とサポート

顧客の利用状況やライフスタイルに基づいて、適切な車両やプランを自動で提案するシステムを導入します。例えば、通勤が主な目的であれば燃費の良いコンパクトカーを、家族旅行が多い顧客にはSUVやミニバンを推奨します。AIやデータ分析を活用し、顧客に合った提案を行うことで、顧客のニーズに応じた柔軟なサービスを提供することが可能です。

理由: 顧客にとって最適な選択を提示することで、利便性や満足度が向上し、長期的な利用やリピート率が上がります。

2. 簡単に柔軟な契約変更手続き

契約期間中にクルマの変更や契約内容の変更が簡単にできるように、オンラインでの手続きやサポート体制を強化します。これにより、ライフスタイルの変化やニーズに応じた柔軟な対応が可能となり、顧客の不満を軽減できます。

理由: 顧客は時間や手続きに煩わされず、サービスをより柔軟に利用できるため、満足度が高まります。また、契約変更が容易であれば、長期的な関係構築にもつながります。

3. 定期的な顧客アンケートとフィードバックの反映

定期的に顧客からフィードバックを収集し、その意見を反映したサービス改善を行う仕組みを導入します。例えば、車両の清潔さやアフターサポートの質、料金プランの透明性などについて、具体的な改善を行い、顧客にその結果を共有することが重要です。

理由: 顧客の声を反映することで、サービスが進化し続け、顧客が「自分の意見が反映されている」と感じることで、エンゲージメントや満足度が高まります。

これらの施策は、顧客のニーズを深く理解し、柔軟かつパーソナライズされた体験を提供することに重点を置いています。結果として、サービスの利便性と顧客体験が向上し、顧客満足度が向上するでしょう。

🌀 思考時間: 8 秒 ↓

お客様の満足度を向上させるために、以下の3つの重要な施策を考案しました。

1. サブスクリプションプランの柔軟性向上

理由: お客様のライフスタイルやニーズは多様であり、時間とともに変化します。契約期間の調整、車種の変更、一時的な利用停止など、柔軟なプランを提供することで、お客様は自分に最適な利用方法を選択できます。これにより、サービスの使い勝手が向上し、満足度の向上につながります。

2. 充実したカスタマーサポートの提供

理由: 車の利用中に問題や疑問が生じた際、迅速かつ確かなサポートが受けられることはお客様にとって大きな安心材料です。24時間対応のサポート窓口や、オンラインチャット、FAQの充実など、サポート体制を強化することで、お客様の信頼と満足度を高めることができます。

3. 付加価値サービスの提供

理由: 基本的な車の提供だけでなく、メンテナンスサービス、保険のパッケージ、ロードサイドアシスタンス、さらには特別なイベントやプロモーションへの招待など、付加価値の高いサービスを提供することで、お客様はより多くのメリットを感じられます。これにより、サービス全体の価値が高まり、顧客ロイヤルティの向上につながります。

これらの施策を選んだ理由は、お客様が求める「柔軟性」「安心感」「付加価値」を直接的に提供できるからです。これらは顧客満足度の主要な要素であり、サービスの差別化にも寄与します。

4

まとめ



まとめ

■ これまでの常識は一旦忘れて、素直な欲求をぶつける

- ・ 「ステップバイステップ」などのテクニックがむしろ悪影響する場合あり
- ・ 「実現したいこと」「必要な情報」にとどめて、論理はAIに任せてみる
- ・ リクエスト数が限られるので「理由も教えて」などアウトプットを増やす工夫が有効

■ GPT-4oとo1-previewの性能差は、もはや素人には分からない領域に

- ・ 解決できる課題が高度になったことで、その判定も高度に
- ・ 現時点ではとりあえず4o、満足いかなければo1を試す流れか
→ただし、4o向けとo1向けではプロンプトを変える必要あり

■ (私見)o1-previewは「プロがよりプロになる」ツールか

- ・ GPT-4oなどは苦手を補ってくれるAIのイメージ
- ・ o1-previewはその道の専門家とも対等にやり取り可能な、プロ向けツールか

Thank you !

