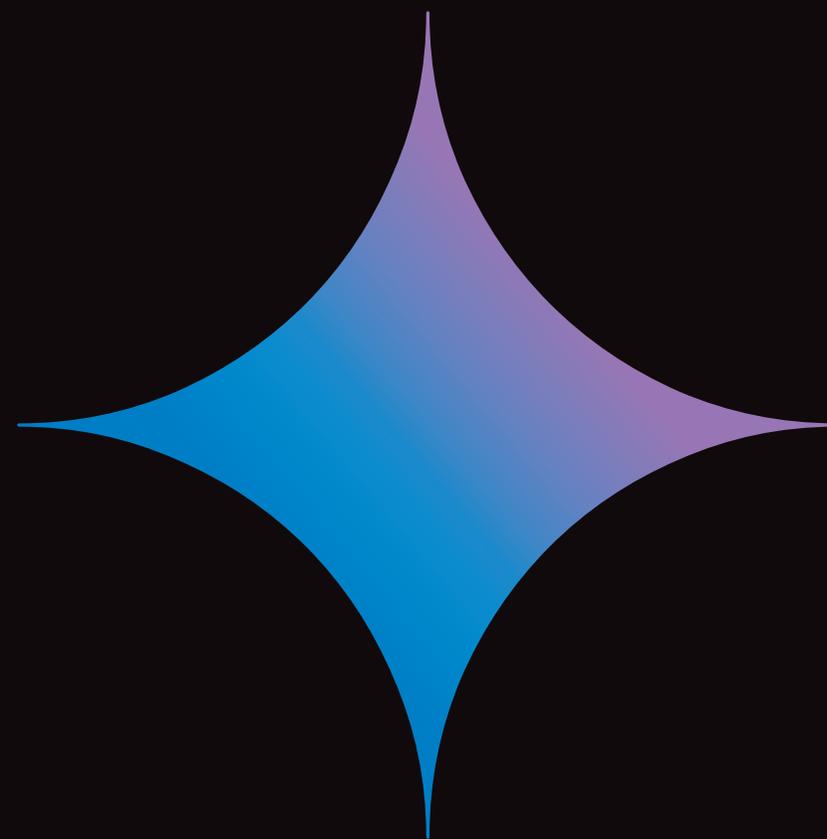


Prompt guide for developers

エンジニアのためのプロンプト活用ガイド



はじめに

エンジニアのみならず、あるいは生成 AI に興味をお持ちのみならず、
こんにちは！
本ガイドブックでは、Google も開発・研究を進めている
最新の生成 AI 技術を、誰でも簡単に活用できるように具体的な
テクニックをご紹介します。

生成 AI の性能を最大限に引き出すには、実は「プロンプト」が重要です。
「プロンプト」とは、生成 AI への指示文のこと。
たとえば、「三角と円を描いてください」という指示を受けたら、
みなさんはどのような図を思い浮かべますか？
人によって解釈が大きく異なる可能性がありますよね。
生成 AI も同様で、明確な指示をしないとどんな図を描けばいいのか
迷ってしまいます。

本書では、Google の最新研究に基づいたプロンプト エンジニアリングの
手法を、具体的な例とともにお伝えします。
明日から使えるテクニック満載なので、生成 AI をもっと活用して、
業務を効率化したり、新しいソリューションを開発したりする
ヒントにしてください！

INDEX

はじめに	01
プロンプト エンジニアリングとは	03
7つのプロンプト テクニック	
1. 複数のタスクに分けて処理する	
2. 処理の順番通りに記述する	
3. DON'T 形式ではなく DO 形式で書く	
4. 複数個の例を提示する	
5. 思考プロセスを例示する	
6. 仕事の内容を本業とする人になりきらせる	
7. 感情を乗せた一言を付け加える	
プロンプト エンジニアリングの実践	14
タスク別プロンプト例	
1. 商品画像からタグ付け	
2-1. テキストのカテゴリ分類(単一テキスト)	
2-2. テキストのカテゴリ分類(複数テキスト)	
3. 商品ページ画像から情報を抽出	
4. 製品紹介文から製品情報を構造化して抜き出し	
5. 動画広告の分析	
プロンプトでは対応できないときの対処法	25
まとめ	28
Google の代表的な生成 AI サービス	29
おわりに	33

プロンプトエンジニアリングとは

生成 AI に期待通りの出力をさせるために、指示文(プロンプト)を設計・改良する技術を指します。特に複雑なタスクでは、効果的なプロンプトエンジニアリングが重要な役割を果たします。また、プロンプトの書き方で性能が大きく変わるという研究結果もあります。

プロンプトに記述する情報の例

役割

あなたは、対話型の AI アシスタントです。与えられた情報を利用して、ユーザーからの質問に正確に回答してください。

質問

日本の野球チームである A は 2023 年に優勝しましたか？

情報

日本のプロ野球についてのニュース記事

出力形式

出力は「はい」か「いいえ」のどちらかです

シンプルなプロンプトに伴う課題

簡易的なプロンプトは、文脈や背景情報が不足しているため、意図しない結果を生成することがあります。

簡易的なプロンプト例

この動画で発生している状況を 3 秒ごとに以下のフォーマットで説明してください。

秒数: [背景]: [状況]



出力例

- 00:00: [Google Pixel のロゴ]: Google Pixel のロゴが表示されます。
- 00:03: [空港]: 若い女性が空港に到着し、携帯電話で写真を撮っています。
- 00:06: [ハワイの風景]: ハワイの青い空とヤシの木が映し出されます。
- 00:09: [空港]: 女性が空港を歩いています。
- 00:12: [ハワイの海岸線]: ハワイの海岸線の美しい景色が広がっています。
- 00:15: [ファーマーズマーケット]: 女性がファーマーズマーケットで携帯電話を使って果物を調べています。
- 00:18: [ランブータン]: 女性がランブータンという果物を試食しています。
- 00:21: [オウム]: 女性が 2 羽のオウムに囲まれて自撮り写真を撮っています。
- 00:24: [ハワイの熱帯植物園]: 女性が植物園で携帯電話を使って花を調べています。

タイムスタンプは 00:00-00:02 形式がいい

俳優の名前を入れてほしい

動画内の会話内容も書いてほしい

説明をより詳しく

生成 AI における良い指示のポイント

指示学習を行ったタスクはうまくできますが、ユーザーが実際に使用するプロンプトは学習に用いたデータとは異なるため、適切なガイドがないとうまく処理できません。仕事を丸投げするのではなく、必要な情報を与えながら適切な指示を出す必要があります。

プロンプトを作成する際は、そもそも人間の視点でも相手に伝わる指示になっているようにした上で、生成 AI に特有の性質を踏まえたプロンプトテクニックを採用していくことが重要です。

人間と共通する事項

- 背景情報・必要な知識を全て説明する
- 明確かつ簡潔に書く
- 作業のプロセスを分解してステップに分ける



生成 AI 特有の事項

- 学習した指示文に沿う
- 地頭の能力を引き出す



改善したプロンプト例

あなたは動画コンテンツ エディターです。以下のタスクに取り組んでください。

[タスク] A. 動画全体を正確なタイムスタンプでシーンに分割します。開始から動画ファイルの正確な終了まで。B. 各シーンについて以下の情報を記述します。B-1. 背景の視覚情報。B-2. 画面上のテキストとロゴ。B-3. シーンで何が起きているか。B-4. B-3 の説明を、背景に見えるものを追加して豊かにします。B-5. 音声情報。

[参考情報] 主人公の名前はアーティスト A です。

[条件] A. 画面上にロゴやテキストがある場合は、単一のシーンとして分割する必要があります。

A. 各シーンの長さは 1~5 秒です。

・良い例: 00:05-00:08, 00:05-00:10。/ 悪い例: 00:05-01:14 タイムスタンプが 1 分以上ジャンプしているため。

B-3, B-4. シーンの説明に音声情報を使用しないでください。画面に見えるものだけに頼ってください。B-4. B-2 で検出されたテキストとロゴを使用してシーンを説明します。画面上に大きなブランドロゴやテキストがある場合は、説明に含める必要があります。B-4. 小説家書いたような、リアルで鮮やかな描写を数文で作成します。

・良い例: 女性が花をスマホで撮影している。花は紫とピンクのグラデーションになっている。
/ 悪い例: 女性が花をスマホで撮影している。B-5. 彼らが何かを言う場合は、「Person1」「彼らが言うこと」Person2「彼らが言うこと」...」のように記述します。そうでない場合は、「なし」と出力します。

[フォーマット命令] 日本語で以下の情報のみを表示します。

4 列のカンマ区切りリストで出力します: タイムスタンプ mm:ss-mm:ss, B-1, B-4, B-5 ヘッダーは次のとおりです: タイムスタンプ, シーン, 説明, 会話



出力例

◆ Timestamp, Scene, Description, Talk

00:00-00:01, Google Pixel logo, Google Pixel のロゴが画面中央に表示される。、

00:01-00:03, 空港, アーティスト A が空港に到着し、スマホで写真を撮っている。背景には「到着」と書かれた青い看板やヤシの木が見える。、アーティスト A「着いたー!ハワイって感じ」

00:03-00:07, 青空とヤシの木, 雲ひとつない青空が広がり、ヤシの木が立ち並ぶ。画面下に「旅も、Google Pixel と。」というテキストとアーティスト A の名前が表示される。、

00:07-00:10, 空港, アーティスト A が空港を出て道路を歩いている。背景には「Ohana by Hawaiian」と書かれた看板が見える。、

00:10-00:11, 道路, アーティスト A が道路を歩いている。足元のみが映っている。、

00:11-00:13, 海と山, 青い海と緑の山が広がる。画面はハワイの雄大な自然を映し出している。、アーティスト A「海だー!」

00:13-00:15, 溶岩台地, 溶岩台地が広がり、奥には山が見える。画面はハワイ島特有の火山地形を映し出している。、

00:15-00:17, マーケット, アーティスト A がマーケットで果物を眺めている。背景には色とりどりの果物や野菜が並んでいる。、

00:17-00:22, ランブータン, アーティスト A がランブータンを手に取り、Google レンズで検索している。画面には「ランブータン」という検索結果が表示される。、アーティスト A「これは何だ?ランブータン。へー、美味しいのかな?」

00:22-00:24, マーケット, アーティスト A がランブータンを食べている。背景にはマーケットの人々が見える。、アーティスト A「あまーい」

00:24-00:26, オウム, アーティスト A が 2 羽のオウムに囲まれ、自撮り写真を撮っている。画面下には Google Pixel のカメラ機能が表示される。、

00:26-00:29, 植物園, アーティスト A が植物園の入り口に立っている。背景には「Hawaii Tropical Botanical Garden」と書かれた看板が見える。、

生成 AI の可能性を最大限に引き出すために、次頁から紹介する [7 つのプロンプト テクニック](#) を学びましょう。

7つのプロンプト テクニック

1. 複数のタスクに分けて処理する 学習したタスク通りに指示を出す

1つのプロンプトで複数のタスクを行わせようとしてうまくいかない場合は、タスクごとにプロンプトを分けてみます。

✖ Before

ポジティブなコメントとネガティブなコメントに分類して。

ポジティブなコメントからこの商品の短い宣伝文を作って。

[コメント]

バッテリーの持ちが良すぎて驚いた！一日中使っても全然余裕。
防水じゃないのが残念。お風呂で使えない。
ケースの種類が少ない。お気に入りのケースが見つからない。
価格が高い。もう少し安いと手が出しやすいのに。
カメラの性能が素晴らしい。特に夜景モードがすごい！
少し重いのが気になる。長時間持つと手が疲れる。
デザインがシンプルでカッコいい。持っているだけで気分が上がる。
動作がサクサクでストレスフリー。ゲームも快適に遊べる。
画面が大きくて見やすい。動画視聴に最適！
指紋認証の精度がいまいち。なかなか認識してくれない。

✔ After (タスクを2つに分割)

ポジティブなコメントとネガティブなコメントに分類して。

[コメント]

バッテリーの持ちが良すぎて驚いた！一日中使っても全然余裕。
防水じゃないのが残念。お風呂で使えない。

コメントからこの商品の短い宣伝文を作って。

[コメント]

バッテリーの持ちが良すぎて驚いた！一日中使っても全然余裕。
カメラの性能が素晴らしい。特に夜景モードがすごい！

タスク1で選択された
ポジティブなコメントを
埋め込む

2. 処理の順番通りに記述する 学習したタスク通りに指示を出す

Deepmindの研究によると、PaLM、Gemini Pro、GPT-3.5、GPT-4 など幅広い LLM で、「自然な推論の流れに沿った指示」を与えない場合、30% 以上精度が落ちる恐れがあると報告されています。

複数のステップからなるタスクでは、プロンプトに記述するステップの順序を変えることで期待する結果が得られる場合があります。

✖ Before

Z は X の 5 倍と Y を足したものです。X は Y の 3 倍です。Y は 100 です。Z は何ですか。

◆ Z は 450 です。

解法:

1. $Y=100$ を X の式に代入します。 $\bigcirc X=Y*3=100*3=300$

2. X と Y の値を Z の式に代入します。 $\bigcirc Z=X*5+Y=300*5+100=1500+100=450$

答え:

● Z=450

1500+100=450と
途中で間違えている

✔ After (処理の順番通りに記載内容の順序を変える)

Y は 100 です。X は Y の 3 倍です。Z は X の 5 倍と Y を足したものです。Z は何ですか。

◆ X は Y の 3 倍なので、 $X=3Y$ となります。

Z は X の 5 倍と Y を足したものですので、 $Z=5X+Y$ となります。

ここで、 $X=3Y$ を $Z=5X+Y$ に代入すると、

$Z=5(3Y)+Y$

$=15Y+Y$

$=16Y$

となります。

よって、Z は Y の 16 倍であることがわかりました。

$Y=100$ とすると、 $Z=16*100=1600$ となります。

結論: Z は 1600 です。

計算手順が明確なので答えが合っている

3. DON'T 形式ではなく DO 形式で書く 学習したタスク通りに指示を出す

否定型の表現の代わりに、肯定型の表現で指示することで期待する結果が得られる場合があります。

✖ Before

Google はどんな会社ですか? Google という単語を使わないでください。

◆ 問題ありません。

「問題ありません」と出力しながら Google と表示している

Google は、世界最大級の検索エンジン、オンライン広告、クラウド コンピューティングなどの事業を展開する、巨大な多国籍テクノロジー企業です。1998 年に設立され、現在では世界中に多くの製品とサービスを提供しています。

✔ After (肯定系の動詞に変換)

Google はどんな会社ですか? Google という単語を避けてください。

◆ 米国を拠点とする多国籍技術会社であり、研究開発、オンライン広告技術、クラウド コンピューティング、コンピューター ソフトウェア、ハードウェア、電子機器に重点を置いています。2023 年現在、親会社である Alphabet Inc. の下で世界最大のテクノロジー企業の 1 つです。

Alphabet Inc. という単語で回避している

4. 複数の例を提示する 地頭の能力を引き出す

期待される入出力の例をプロンプト内で大量に提示する手法を **Many-shot prompting** と呼びます。「###, <<>, #入力, 例」など、データを示す記号は自由ですが、一貫した形式で記述しない場合、性能が下がることが知られています。

✖ Before

次の文章をポジティブ、ネガティブ、中立に分類分けしなさい。
ただし、ポジティブは 1、ネガティブは 2、中立は 3 と出力しなさい。

この映画は退屈だったけど、最後のシーンは感動した。

◆ 2

この映画は退屈だったけど、最後のシーンは感動した。
この文章は、映画に対して"2 (ネガティブ)"と判定します。

理由:

- ・前半部分の「退屈だった」という表現は、映画全体に対してネガティブな評価をしています。
- ・後半部分の「最後のシーンは感動した」という表現は、映画の最後のシーンに対してポジティブな評価をしています。

✔ After (「例: 」という記号で何のための情報かを明記)

次の文章をポジティブ、ネガティブ、中立に分類分けしなさい。
ただし、ポジティブは 1、ネガティブは 2、中立は 3 と出力しなさい。

例: この本は面白かったけど、最後の展開はありきたりだった。//3 と出力。

例: この曲は好きだけど、歌詞は意味が分からない。//3 と出力。

例: この映画は主演の演技が最高で最後まで見応えがあった。//1 と出力。

この映画は退屈だったけど、最後のシーンは感動した。

◆ 3

この映画は退屈だったけど、最後のシーンは感動した。
この文章は、"3 (中立)"と判定します。

ポジティブとネガティブ両方の要素を持つ際は中立とすることを例から理解した

理由:

- ・前半部分の「退屈だった」という表現は、映画全体に対してネガティブな評価をしています。
- ・後半部分の「最後のシーンは感動した」という表現は、映画の最後のシーンに対してポジティブな評価をしています。
- ・この文章は、映画に対してネガティブとポジティブの両方の評価を含めたため、中立と判定します。

5. 思考プロセスを例示する 地頭の能力を引き出す

複雑な思考プロセスが必要な問題については、**Few-shot Prompting** で提示する例において、思考プロセスをあわせて示すことで性能が向上します。若手コンサルタントが過去の事例から手法を学び、同じ方法でアウトプットするのに例えられます。

Chain of Thought Prompting と呼ばれる手法です。

✖ Before

100000000000+34321*10^2 は何ですか。日本語で答えなさい。

◆ 1100000000000+34321*10^2=100003432100 になります。

✔ After

100000000000+34321*10^2 は何ですか。日本語で答えなさい。

例: $1+1*10^2$ は何ですか。

出力例: 10^2 は 100 です。 $1+100$ は 101 なので答えは 101 です。

例: $10000000+45*10^2$ は何ですか。

出力例: 10^2 は 100 です。 $45*100$ は 4500 なので、 $10000000+4500$ は 10004500 です。

◆ $100000000000+34321*10^2$ を計算するために、以下のように考えてみましょう。

1. 10^2 は 10 の 2 乗なので、計算すると 100 になります。

2. 34321 に 100 をかけた結果は次のようになります。

$$34321*100=3432100$$

3. これを 100000000000 に加算します。

$$100000000000+3432100=100003432100$$

したがって、 $100000000000+34321*10^2$ は 100003432100 になります。

6. 仕事の内容を本業とする人になりきらせる 地頭の能力を引き出す

論理的な作業や IT に関連する作業をさせたい場合に、例えば「あなたは Google の Staff Engineer です。」と付け加えると性能が上がるという結果が知られています。

✖ Before

100000000000+34321*10^2 は何ですか。日本語で答えなさい。

◆ $100000000000+34321*10^2=100003432100$ になります。

✔ After

あなたは Google の Staff Engineer です。

100000000000+34321*10^2 は何ですか。日本語で答えなさい。

◆ $100000000000+34321*10^2=100000000000+3432100=100003432100$
Google の Staff Engineer である私は、この計算を正確に行うことができました。

7. 感情を乗せた一言を付け加える 地頭の能力を引き出す

「これは私のキャリアにとって大事です。」「成長の機会だと考えて頑張ってください。」「頑張れ、君ならできる!」などの言葉で性能が上がる場合があることが知られています。

✖ Before

100000000000+34321*10^2 は何ですか。日本語で答えなさい。

◆ $100000000000+34321*10^2=100003432100$ になります。

✔ After

100000000000+34321*10^2 は何ですか。日本語で答えなさい。君ならできる頑張れ!

◆ $100000000000+34321*10^2=100000000000+3432100=100003432100$

参考

機械学習モデルから見たプロンプトの役割

- 大規模言語モデルは、巨大なニューラル ネットワークに対して、実在する膨大な文書を学習させた結果、ネットワーク内部に「さまざまな人格」「さまざまな知識」「さまざまな処理能力」が同時に埋め込まれています。
- プロンプトに入力した指示・データによって、「ネットワークのどの部分を利用して結果を生成するか」が変わります。
- ペルソナの指定や感情を乗せたプロンプトによって、ネットワーク内の利用する部分が変化して、出力結果が変わるものと考えられます。

プロンプト エンジニアリングの実践

構造化プロンプトの活用

多くの場合、プロンプトに記述すべき項目は決まっているので、次のようなテンプレートを用意しておくといでしょう。多くの LLM で平易かつ明確な英語のプロンプトの方が性能が良くなるという結果が報告されていますので、可能であれば英語で書くことを推奨します。

```
You are a ----- . Work on the following tasks.  
  
[task]  
A. -----  
B. -----  
  
[reference information]  
----  
  
[condition]  
A. -----  
B. -----  
  
[format instruction]  
Output in Japanese. -----  
  
[context]  
-----
```

仕事の内容を本業とする人になりきらせて上で、情報をジャンルごとにまとめてラベルを振る

プロンプト実装のステップ

- プロンプトの構成パターンは前頁までで紹介した例だけとは限りませんが、基本的には、「タスクをステップに分けて指示する部分」と「LLM が知らない情報の部分」と「詳細な条件を指定する部分」をどのように書くかを考えていきます。
- 始めから複雑な指定をするのではなく、「シンプルな内容から始めて、出力結果を見ながら、徐々に指示を追加」していきます。
- 「シンプルなプロトタイプで基本機能を実装した後に、テストしながらより複雑な機能を追加していく」というプログラム開発と同じ考え方になります。
- ただし、どのような指示がどんな変化を引き起こすかは予測困難なので、試行錯誤が必要です。

タスク別プロンプト例

1. 商品画像からタグ付け

商品画像のファイルからそれを表現する SNS 等で使用するタグを生成します。

You are a content editor of an EC web site. Work on the following tasks.

[task]
A. Generate tags that represent products in the picture.

[condition]
A. Avoid tags that represent general category such as drink, food, toy. A. Use tags only for concrete things that exist in the picture. A. Include tags that describe product features.

[format instruction]
In Japanese. In plain text, no markdown. No header. Show tags in a comma separated list.



具体的な条件を指定することで意図した通りの出力が期待できる

◆ クロワッサン, コーヒー, マグカップ, 木のテーブル, 皿, フォーク, ナイフ

2-1. テキストのカテゴリ分類(単一テキスト)

与えられたテキストのポジネガ判定を理由を添えて行います。

You are a psychologist specialized in text analysis. Work on the following tasks.

[task]
A. Classify the given text into two categories: positive or negative.
B. Give a reason in a single short sentence.

[condition]
A, B. Give an answer consistent with examples.

[format instruction]
In Japanese. In plain text. No markdown. No header. Output in a comma separated list with columns: category, reason

[examples]
- text: この傘、すごく軽いいし、持ち運びに便利そうでいいですね！特に折り畳み式なのが気に入りました。でも、デザインがちょっと地味かな...もう少し明るい色や柄があったらよかったかも。とは言え、機能性は抜群そうなので、旅行用に使います。 / category: ポジティブ / reason: 最終的に機能性は抜群で旅行用に使うと言っているから。
- text: サイズが全く合いませんでした。残念です。デザインは素敵だったのですが、サイズ表記が正確ではないようです。返品です。 / category: ネガティブ / reason: サイズが合わなくて残念と言っており、最終的に返品と言っているから。

[text]
デザインがすごく可愛くて一目惚れして買ったんだけど、ちょっとサイズが小さかったみたいで長時間履いていると足が痛くなっちゃうのが残念。ワンサイズ大きめを買えばよかったかな。でもデザインは本当に気に入ってるから、近場のお出かけの時に履こうと思います。

理由をあわせて説明させることで合理性のある判断をしているか確認する

例を参照して一貫した判断基準を適用させる

例を増やすことで、より評価基準に即した結果が得られるようになる

思考プロセスを示すことで暗黙的な判断基準を学ばせる

◆ ポジティブ, デザインを気に入っており、近場のお出かけの時に履くと言っているから。

2-2. テキストのカテゴリ分類(複数テキスト)

複数の文章を文章ごとに番号を割り振り、それぞれの文章がどの感情を表現しているか分類します。

You are a psychologist specialized in text analysis.
Work on the following tasks.

[task]

A. Do the followings for each text given in the [text] section.

A1. Give a sequential number starting from 1.

A2. Select a part of the text that represents the emotion.

A3. Select one of the categories in the [category] section.

[category]

喜び、悲しみ、期待、驚き、怒り、恐れ、嫌悪、信頼

[condition]

A. The [text] section contains one text for each line.

A3. If there's no category that fits the given text, give "その他(xxxx)" as its category where xxxx is a word of emotion that represents that text.

[format instruction]

In Japanese. In plain text. No markdown. No header.

Output in a comma separated list with columns: sequential number(A1), category(A3), "text", "selected part(A2)".

Use double quote for "text", "selected part(A2)".

[text]

ぼけっとしてたらこんな時間。チャリあるから食べに出たいのに...

今日の月も白くて明るい。昨日より雲が少なくてキレイだな〜と立ち止まる帰り道。チャリなし生活も悪くない。

早寝するつもりが飲み物がなくなりコンビニへ。ん、今日、風が涼しいな。眠い、眠れない。表情筋が衰えてきてる。まずいな...やばい。(後略)

ステップに分けてやることを指示する

- ◆ 1. その他(焦り), "ぼけっとしてたらこんな時間。チャリあるから食べに出たいのに...", "ぼけっとしてたらこんな時間"
- 2. 喜び, "今日の月も白くて明るい。昨日より雲が少なくてキレイだな〜と立ち止まる帰り道。チャリなし生活も悪くない。", "キレイだな〜"
- 3. その他(のんびり), "早寝するつもりが飲み物がなくなりコンビニへ。ん、今日、風が涼しいな。", "風が涼しいな"
- 4. その他(焦り), "眠い、眠れない。", "眠い、眠れない"
- 5. 怒り, "ただいま〜って新体操してるやん!外食する気満々で家に何も無いのに!テレビから離れられない...!", "外食する気満々で家に何も無いのに!"
- 6. 悲しみ, "表情筋が衰えてきてる。まずいな...", "表情筋が衰えてきてる"
- 7. その他(空腹), "おなかすいた...夜ご飯仲間募集。", "おなかすいた"
- 8. その他(混乱), "よく寝た。暴走バスの夢見た。からだいたい。", "暴走バスの夢見た"

3. 商品ページ画像から情報を抽出

電子書籍の商品画像ページから複数形式(電子書籍、ソフトカバー)の情報を取得し、著者かつ監修者の情報を正しく処理します。

You are a product analyst of a book EC web site.
Work on the following tasks.

[task]

A. Detect person indicated as (著, 監修) who is both an author and a supervisor.

B. Detect person indicated as (著) who is an author.

C. Detect person indicated as (監修) who is a supervisor.

D. Identify the following information of the book.

D1. Book title

D2. Format

D3. Authors

D4. Supervisors

D5. Price

D6. Description

[condition]

D2. The format is "電子書籍" or "ソフトカバー" or "ハードカバー".

D2. The book may have multiple format. In that case, generate multiple records.

D3, D4. A person could be both an author and a supervisor. In that case, that person should be in both D3 and D4.

[format instruction]

In Japanese. In plain text, no markdown. No header. Show result in a comma separated list with columns that correspond to "D1", "D2", "D3", "D4", "D5", "D6" with double quote.

●前提となる作業を分けて指示する

●作業の前提条件を明示的に与える

The screenshot shows a book page with the following details:

- おすすめ** (Recommended)
- エンタープライズのための Google Cloud クラウドを活用したシステムの構築と運用**
- 遠山 雄二 (著, 監修), 矢口 悟志 (著), 小野 友也 (著), 渡邊 誠 (著), 岩成 祐樹 (著), 久保 智夫 (著), 村上 大河 (著), 星 美鈴 (著), 中井 悦司 (監修), 佐藤 聖規 (監修)
- クラウド コンピューティングが普及する中で、エンタープライズ システムもまた、オンプレミスからクラウドへの円滑かつ効果的な移行が求められています。本書は、エンタープライズ システムを構築・運用するエンジニアのために、Google Cloud の具体的なユースケースや設計ポイントを Google Cloud のエンジニア陣が徹底的に解説する一冊です。
- 4.3 ★★★★★ ☆ 38 個の評価
- 本のまとめ買いで最大 15% ポイント還元 3/20 まで 今すぐエントリー

Category	Price	Quantity
電子書籍 - すぐに読めます	¥1,815 (20pt 還元)	- 0 +
単行本(ソフトカバー)	¥3,630 (200pt 還元)	- 0 +

- Add to Cart** (Add to Wishlist)
- Details:** 2022/1/18 発刊
無料お届け日時の指定もできます
詳細を見る

◆ "エンタープライズのための Google Cloud クラウドを活用したシステムの構築と運用", "電子書籍", "遠山雄二, 矢口悟志, 小野友也, 渡邊誠, 岩成祐樹, 久保智夫, 村上大河, 星美鈴", "遠山雄二, 中井悦司, 佐藤聖規", "¥1,815", "クラウド コンピューティングが普及する中で、エンタープライズ システムもまた、オンプレミスからクラウドへの円滑かつ効果的な移行が求められています。本書は、エンタープライズ システムを構築・運用するエンジニアのために、Google Cloud の具体的なユースケースや設計ポイントを Google Cloud のエンジニア陣が徹底的に解説する一冊です。" "エンタープライズのための Google Cloud クラウドを活用したシステムの構築と運用", "ソフトカバー", "遠山雄二, 矢口悟志, 小野友也, 渡邊誠, 岩成祐樹, 久保智夫, 村上大河, 星美鈴", "遠山雄二, 中井悦司, 佐藤聖規", "¥3,630", "クラウド コンピューティングが普及する中で、エンタープライズ システムもまた、オンプレミスからクラウドへの円滑かつ効果的な移行が求められています。本書は、エンタープライズ システムを構築・運用するエンジニアのために、Google Cloud の具体的なユースケースや設計ポイントを Google Cloud のエンジニア陣が徹底的に解説する一冊です。"

4. 製品紹介文から製品情報を構造化して抜き出し

製品の紹介文から製品名・製造メーカーを抽出します。

You are a professional curator of a video game museum. Work on the following tasks.

[task]
A. Extract service name and service description from the given text.

[condition]
A. Use the reference information to infer the output.

[format instruction]
Use this JSON schema:

```
{ "type": "object",
  "properties": {
    "game information": {
      "type": "object",
      "properties": {
        "service name": { "type": "string" },
        "service description": { "type": "string" }
      }
    }
  }
}
```

[reference information]
The following is a comma separated list of game title and game maker.

```
-----
service name, service description
-----
Gmail, メール
Drive, ファイル管理
Calendar, スケジュール管理
Meet, ビデオ会議
Chat, チャット
-----
```

● JSON schema 言語でフォーマットを指定

● 既知の参考情報を提供

⋮

Docs, ドキュメント管理
 Sheets, 表計算
 Slides, プレゼンテーション
 Forms, アンケート
 Sites, ウェブページ作成
 Keep, メモ
 Apps Script, スクリプティングツール

[text]
大容量ストレージをチームで活用

➤

◆

```
```json
{
 "service information": {
 "service name": "Drive",
 "service description": "ファイル管理"
 }
}
...
```
```

5. 動画広告の分析

動画広告からブランドロゴの有無やカット数などを要素に分けて分析します。

You are a video content analyst specialized in advertisement.

Work on the following tasks.

[task]

A. Detect texts on the screen.

- A1. Check if the text detected on A is a brand logo.
- A2. Check if the text detected on A is shown in a large size.

B. Detect things shown at the beginning of the movie.

- B1. Check if the thing detected on B is a close-up shot.

C. Check how many times camera angles or scenes change.

- C1. Calculate the number of scenes that is the number of changes plus 1.

[condition]

A, B, C. Analyze on first 5 seconds only, starting from 00:00 through 00:04. Avoid using information after 00:05.

[format instruction]

In Japanese. In plain text. No markdown.

Output in the following format.

- A1. Yes or No : explanation
- A2. Yes or No : explanation
- B1. Yes or No : explanation
- C1. Number : explanation in order of timestamp. A comma separated list with columns: timestamp, scene



具体的な手順をステップに分けて指示

冒頭 5 秒以外の情報は参照しないように指示

- ◆ A1. はい: Google Pixel 8 のロゴが表示されています。
- A2. はい: 画面の大部分を占めるようにロゴが表示されています。
- B1. はい: 女性の顔が画面いっぱいに映るクローズアップ ショットです。
- C1. 3:
00:00, 女性の顔のクローズアップ
00:02, 駅の券売機と通行人の様子
00:03, 券売機の画面のクローズアップ

良い指示の具体例のプロンプトはこちらから確認できます

1. 商品画像からタグ付け



2-1. テキストのカテゴリ分類(単一テキスト)



2-2. テキストのカテゴリ分類(複数テキスト)



3. 商品ページ画像から情報を抽出



4. 製品紹介文から製品情報を構造化して抜き出し



5. 動画広告の分析



* 生成 AI の特性上、本ガイドブックに掲載しているレスポンス内容と異なる出力になる場合があります。あらかじめ、ご了承ください。

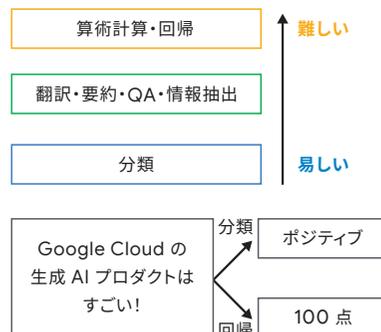
プロンプトでは 対応できないときの対処法

LLM が得意なタスクに変更する

生成 AI はさまざまなタスクをこなせるように進化していますが、算術計算など苦手な分野もあります。

レビュー コメントを評価する際に、点数で評価する(回帰問題として解く)代わりに、5段階評価に置き換えて分類問題として解くこともできます。

生成 AI が得意とする方法に問題を置き換えるという発想も時には有効です。



API サービスで対応する

タスクを複数のステップに分けて処理する際に、一部のステップを他の API サービスに依頼して処理するという方法も考えられます。この際、「どの API サービスを使うべきか」「API サービスへの入力値は何にするのか」という判断を生成 AI で行います。このような処理を行う Function calling と呼ばれる機能を持つ生成 AI もあります。

生成 AI によるエージェントを作成して対応

API サービスとの連携をさらに高度化すると、「何をすべきかを考える(推論)」「API サービスでそれを実行する(行動)」という処理を多段階で繰り返して最終結果を生み出す「生成 AI によるエージェント」を作ることができます。これは、ReAct (Reasoning and Acting) と呼ばれる手法です。

エージェントを擬似的に実現する

エージェントの仕組みを実感するために、特別なツールを使用せずに、手作業でエージェントを擬似的に実現する例を示します。

ステップ 1: プロンプトにツールの情報を付与した上で、質問に答えるように要求します。

依頼事項
質問に答えてください。
ツールを使用する必要がある場合は、ツール名と入力データを答えてください。

[ツール]
ツール名: 天気 API
→ 機能: {location: '都市名'} を入力すると、指定した都市の現在の天気を得られます。

[質問]
東京の現在の天気を教えてください。

あなたの回答 1

◆ [ツール] ツール名: 天気 API 入力データ: {location: '東京'} 生成 AI は使用するツールと入力データを答える

ステップ 2: これまでのやりとりに、LLM が指定したツールを呼び出した結果を追加したプロンプトを改めて入力します。

依頼事項
質問に答えてください。
ツールを使用する必要がある場合は、ツール名と入力データを答えてください。

[ツール]
ツール名: 天気 API
→ 機能: {location: '都市名'} を入力すると、指定した都市の現在の天気を得られます。

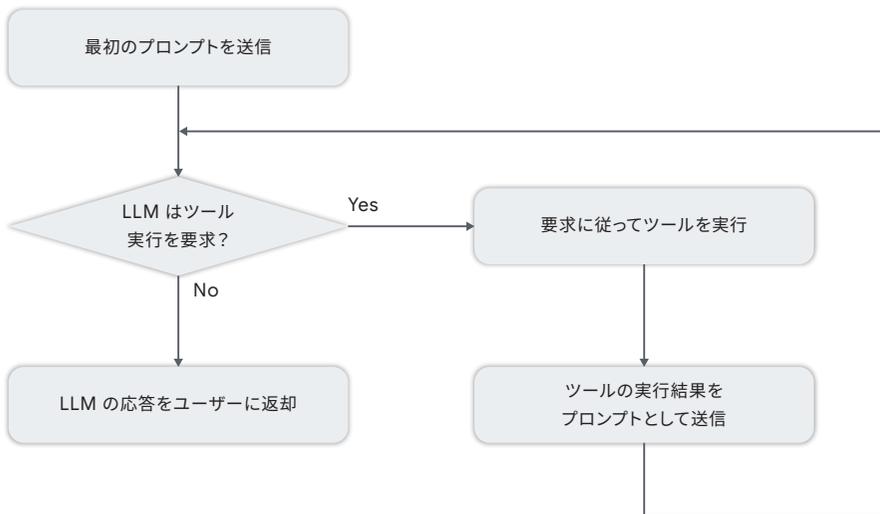
[質問]
東京の現在の天気を教えてください。

あなたの回答 1
[ツール] ツール名: 天気 API 入力データ: {location: '東京'}
ツールの応答
{location: '東京', weather: '晴れ', 気温: '32 度'}
あなたの回答 2

◆ 東京の現在の天気は晴れで、気温は 32 度です。

Gemini API の Function Calling 機能で LLM Agent を実装する

Gemini API の Function calling 機能を利用すると、以下の処理を実装したプログラムを簡単に作成できます。複数のツールを組み合わせさせて使わせることも可能です。



まとめ

3つのポイント

- ✓ プロンプト エンジニアリングとは、生成 AI / LLM に期待通りの出力をさせるための指示文設計・改良技術。プロンプトの書き方で性能が大きく変わる。
- ✓ プロンプトを書く際は定型フォーマットから始めて、性能向上のためにさまざまな手法を追加して試行錯誤するのが大事。
- ✓ 生成 AI に向いている問題設計をしたり、Function Calling で API 連携をして機能を拡張するなど、「生成 AI にできること」を理解した設計も重要。

7つのプロンプト テクニック

- | | |
|------------------------|---------------------------|
| 1. 複数のタスクに分けて処理する | ✓ 1つのタスクに特化して指示 |
| 2. 処理の順番通りに記述する | ✓ 複数ステップのタスクでは自然な推論の順序で指示 |
| 3. DON'T 形式ではなく DO 形式で | ✓ 肯定系の動詞を使う |

Many-shot Prompting

- | | |
|---------------|---------------------|
| 4. 複数個の例を提示する | ✓ 期待される入出力の例を複数提示する |
|---------------|---------------------|

Chain of Thought Prompting

- | | |
|----------------|-----------------|
| 5. 思考プロセスを例示する | ✓ 例の中で思考プロセスを示す |
|----------------|-----------------|

- | | |
|------------------|----------------------|
| 6. 本業とする人になりきらせる | ✓ タスクに適した専門家の役割を設定する |
|------------------|----------------------|

- | | |
|-------------------|---------------------|
| 7. 感情を乗せた一言を付け加える | ✓ モチベーションを高める言葉を添える |
|-------------------|---------------------|

Google の 代表的な生成 AI サービス

| コンシューマ向け
生成 AI サービス | 企業向け生成 AI サービス | |
|-----------------------------|---|---|
| | Google Workspace
ユーザー向け | Google Cloud
ユーザー向け |
| Gemini
(旧称: Bard) | Gemini for
Google Workspace
(旧称: Duet AI for
Google Workspace) | Gemini for
Google Cloud
(旧称: Duet AI for
Google Cloud) |
| Google AI Studio | Gemini for AppSheet
(旧称: Duet AI for
AppSheet) | Vertex AI

Vertex AI Studio

Vertex AI Agent
Builder |

コンシューマ向け生成 AI サービス

Gemini

Google DeepMind によって開発されたマルチモーダル大規模言語モデル。Gemini は、データセンターからモバイル デバイスまで、あらゆる環境で効率的に実行が可能。

Google AI Studio

Gemini を使った生成 AI アプリケーションを開発したりプロトタイピングができるサービス。応答の出力の制御や API キー発行も可能。

企業向け生成 AI サービス - Google Workspace ユーザー向け

Gemini for Google Workspace

Google Workspace で常時有効な AI アシスタント。Gemini をトレンドの把握、情報の統合、ビジネス機会の特定に役立つリサーチ アナリシスなどさまざまな用途で利用が可能。データはユーザーや組織が自身で管理できるため、ビジネス利用にも安全で有用。

Gemini for AppSheet

既存のシートから簡単にアプリを作れるノーコード開発ツール。Google Chat を介して生成 AI の力によって業務のサポートが可能。

企業向け生成 AI サービス - Google Cloud ユーザー向け

Gemini for Google Cloud

開発者、Google Cloud サービス、およびアプリケーション向けの新世代の AI アシスタント。Gemini Code Assist、Gemini Cloud Assist、Gemini in Security、Gemini in Threat Intelligence、Gemini in BigQuery、Gemini in Looker、Gemini in Databases など開発者からビジネスアナリスト、セキュリティ担当者など多くのビジネスパーソンのパフォーマンスの向上、効率化が実現可能。

Vertex AI

ML モデルと AI アプリケーションのトレーニングとデプロイを行い、AI を活用したアプリケーションで使用する大規模言語モデル (LLM) をカスタマイズできる機械学習 (ML) プラットフォーム。機械学習全般とカスタムモデルの構築が可能。

Vertex AI Studio

開発するアプリケーションを Google Cloud 上の他サービスと、シームレスに連携する場合や、アプリケーションとして公開したり生成 AI モデルのプロトタイピングやカスタマイズを行うことができるプラットフォーム。

Vertex AI Agent Builder

Google の生成 AI 技術を活用したアプリケーションをローコードで作成できるプロダクト。

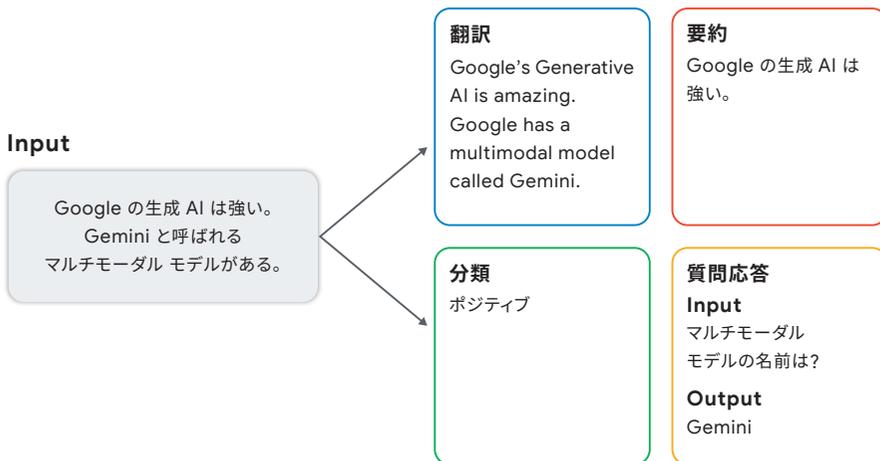
参考

大規模言語モデル (LLM) とは

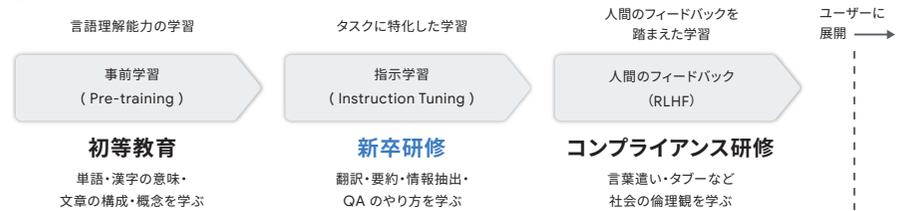
生成 AI を構成する大規模言語モデル (LLM) は、本質的には大規模なニューラル ネットワークであり、「次の単語を予測する」という処理を繰り返し行うことで、プロンプトに続く応答テキストを生成しています。つまり、LLM は膨大な数の「人間が作った文書 (= 論理的整合性を持った文書)」で学習した結果、論理的整合性を持った文書を生成する能力を獲得したといえます。言い換えると、LLM に期待する仕事をさせるには論理的な指示が必要であり、曖昧な入力からは曖昧な出力が得られることとなります。

生成 AI の学習プロセスについて

生成 AI は、従来の機械学習では個別のモデルを使っていたさまざまなタスクが実行できます。ただし、何でも自由にできるわけではありません。それは、生成 AI の学習プロセスに理由があります。生成 AI は、タスクに特化しない事前学習を行った後、要約・分類・質問回答など、特定のタスクに特化した学習データでファインチューニング (指示学習) を行います。そのため、指示学習を行ったタスクは比較的うまく処理します。



[生成 AI の学習を人間で例えると]



In-Context Learning

In-Context Learning (ICL) とは、LLM がプロンプトで示されたインプット/アウトプットの例からタスクを学習する能力のことです。例を与えるだけで、推論の精度を上げることができます。いくつかの例示をするのを Few-shot と呼び、大量の例示をするのを Many-shot と呼びます。また、例示なしの場合を区別して Zero-shot Prompting と呼ぶことがあります。

Chain of Thought Prompting

Chain of Thought Prompting (CoT プロンプティング) は、In-Context Learning において思考プロセスの例を示し、LLM にも同様の思考を促すことで、LLM がより正確な回答を生成しやすくなる手法です。たとえば、「りんごが 5 個、みかんが 3 個あります。合わせて何個ですか?」という問題に対して、「りんごが 5 個とみかんが 3 個なので、足し合わせて 5+3=8 個です。」のように、思考過程を含めた例を LLM に提示します。

Zero-shot Prompting

例を提示せずに、言葉だけで指示する。

Prompt 例

みかんはどんな果物ですか?

Few-shot Prompting

少数の例を提示して、その例に基づいたものを生成する。

りんごは赤い果物です。バナナは黄色い果物です。レモンは黄色い果物です。では、みかんは何色の果物ですか?

Many-shot Prompting

大量の例を提示して、その例に似たものを生成する。

りんごは赤い果物です。バナナは黄色い果物です。レモンは黄色い果物です。ぶどうは紫色の果物です。メロンは緑色の果物です。イチゴは赤い果物です。では、みかんは何色の果物ですか?

Chain of Thought Prompting

思考プロセスの例を提示する。

以下のようなステップに従って、みかんが何色かを推測してください。

質問: りんごは何色かを推測してください。

ステップ 1: りんごはバラ科に属します。

ステップ 2: イチゴやサクランボなど、バラ科の果物は熟すと赤やピンク色になることが多いです。

結論: りんごも赤色だと予測できます。

おわりに

本書で紹介したプロンプト テクニックは、
生成 AI を活用するための第一歩です。
実際に手を動かしてプロンプトを試してみることが、
さらなる活用への近道となります。

たとえば、動画分析で思ったような結果が得られなくても、
プロンプトを工夫することで、「ブランドロゴの検出」や
「シーン構成の分析」など、より高度な分析ができるようになります。
諦めずにプロンプトを改善することで、文書作成、コード生成、データ分析
などさまざまな業務でニーズを満たすような回答が得られるはずです。

私たち Google は、みなさまの挑戦をサポートしていきます。
生成 AI を活用して、新しい価値を生み出していきますよう！

Gemini をあなたの企業でも

今すぐ gemini.google.com で
試してみてください

Gemini の
企業活用に関する
お問い合わせはこちら



cloud.google.com/contact

Gemini 公式 note で
情報発信中
今すぐチェック



note.com/google_gemini/

Google Cloud

cloud.google.com

2025 年 2 月版