

※出題範囲は以下の内容を含みますが、これらに限定されるものではありません。

## 出題範囲（参考訳）

### 生成AIの手法と方法論

#### 生成AIを定義する

生成AIと、予測系AI、識別系AI、分析系AI、統計AIを比較・対比する

生成AIと検索エンジンを比較・対比する

拡散モデル、トランスフォーマーモデル、敵対的生成ネットワーク（GAN）、変分オートエンコーダー（VAE）の基本を理解する

#### 生成AIの基本的な出力プロセスを説明する

モデルごとにトレーニング方法が違うことを理解する

- テキスト生成モデルの例：OpenAI GPTx, Google Gemini, Anthropic Claude, Meta LLaMA

- 画像生成モデルの例：DALL-E, Adobe Firefly

大規模言語モデル（LLM）が効果的に機能するためには、大量のトレーニングデータが必要となる

LLMのトレーニングに使用されるデータセットは膨大なため、なかには様々な意見や観点も含まれる

画像生成モデルは、手動でタグ付けしたテキストと画像のペアを使ってトレーニングされる

モデルのトレーニングには大量の電力と強力な画像処理装置（GPU）が必要となる

モデルの種類：拡散モデル、トランスフォーマーモデル、敵対的生成ネットワーク（GAN）、変分オートエンコーダー（VAE）

キーワード：ニューラルネットワーク、畳み込みニューラルネットワーク、トークン化、拡散、ノイズ生成、改良アルゴリズム、ハイパーパラメーター、データセット

#### 生成AIによる入出力の種類を特定する

複数の入力から1つの出力を得ることができる

- 入力の例：テキスト、音声、動画、画像
- 出力の種類の例：生成テキスト、生成動画、生成画像、生成音声

ツールごとに異なる種類の入力ができる

#### 生成AIモデルが個人のタスクに合わせてカスタマイズできることを理解する

自分の代わりにタスクを行ってくれる自己完結型アプリ

- 例：Custom GPT, Google Gems, Microsoft Copilots

#### タスクに合わせて適切なツールを選択する

ツール：Microsoft Copilot, Google Gemini, MetaGPT, Adobe Express, Canva, Open AI ChatGPT, Claude, Microsoft Azure AI Studio, Stable Diffusion

ツール選択時の考慮事項：目的と機能性、使いやすさ、コスト、更新とサポート、個人情報保護、セキュリティ、品質、カスタマイズ性、出力制御のパラメーター

#### 生成AIの限界を説明する

出力が信頼できるとは限らない

出力にバイアス、誤情報、ハリシネーションが含まれる可能性がある

処理能力とデータへのアクセス（通常はインターネット）が必要となる

プライバシー設定を有効化しないと、AIとのやり取りがトレーニングに使用される

利用方法についての世界共通基準が存在しない

一貫性の限界（例：2つの時計が異なる時刻を示す）

急速な変革により、以前作成したものがすぐに時代遅れになる可能性がある

## 出題範囲（参考訳）

### 基本的なプロンプトエンジニアリング

#### テキスト情報を引き出すのに適したプロンプトを特定する

コンテンツを収集する

要約する

コンテンツを作成する、アイデアを創出する

#### コンテンツの変換に適したプロンプトを特定する

要件に合わせてコンテンツの書式を再設定する

ドキュメントの編集と校正を行う

コンテンツを視覚化する

コンテンツを別の種類のメディアに変換する

コンテンツを別の言語に翻訳する

学習と理解促進のためにコンテンツを個人に合わせてカスタマイズする

#### 画像の生成や変換に適したプロンプトを特定する

特定の目的のために画像を制作する

芸術的なアイデアを探求する

画像を変換する

画像の内容を説明する

#### 動画の生成や変換に適したプロンプトを特定する

画像に動きを追加する

画像間の補間を行う

白黒映画をカラー化する

プロンプトから動画を生成する

スクリプトを読むアバターを生成する

動画にオブジェクトを追加する、削除する

字幕を自動表示する

### プロンプトの改善

#### 初期プロンプトとその出力をもとに、より的確な出力を引き出せるプロンプトへの改善方法を評価する

コンテンツ

- 過不足のないプロンプトを作成する
- 明確で省略のないプロンプトを作成する
- AIが自分の意図を「理解してくれるだろう」と思い込まない

スタイル

- 出力のスタイルやトーンに関する情報を含める
- スタイルガイドを含める

ペルソナ

- AIに特定の人格または役割を与える

コンテキスト

- AIに指示を出す際は、背景情報を与える必要がある。AIは機械であり、自ら文脈を理解することはできない

## 出題範囲（参考訳）

初期プロンプトとその出力をもとに、より的確な出力を引き出せる追加の入力情報を特定する

例（few-shotプロンプト）

翻訳用の用語集

テンプレート

調査に用いる文献

同一スレッド内の過去の対話内容

一般的なプロンプト手法を認識する

Zero-shot, few-shot, chain-of-thought, self-consistency（自己整合性）、generate knowledge（知識生成）、プロンプトチェーン

リバースプロンプト手法を使って結果を得る

AIの出力をもとに、その正確性を検証する方法を説明する

歴史的事実

現在の事実

数値データ

倫理、法律、社会的影響

生成AIの出力にバイアスが含まれる可能性を特定する

AIはトレーニングデータに含まれるバイアスを出力に反映する可能性がある

異なるバイアスが、モデルごとに存在する可能性がある

モデルの作成者がガードレールを加えることで、バイアスが生じる可能性がある

- 一部のツール（例：AzureOpenAI）では、追加されたガードレールのオン・オフを切り替えることができる

プロンプトによるバイアスが生じる可能性がある

一般的なバイアスには、性別、人種、障がい、年齢、宗教、文化、言語、国籍、経済状況などがある

生成AIはバイアスを拡散する手段になり得る

生成AIの使用に伴う潜在的な法的影響を特定する

知的財産権を尊重する

- 法制度が流動的な現在、最善策はAIを使用せず、他人の作品を許可なく使用しないことである

- 一部のモデルでは著作権保護されたデータがトレーニングに使用された

生成されたコンテンツの不適切な使用に伴う法的影響を特定する

透明性の確保（職場でAIを使用する際はプロセスを記録する）

データプライバシーの重要性を説明する

個人情報や企業の機密情報がトレーニングに使用されることがある

個人を特定できる情報（PII）が生成AIに使用されると、個人情報の盗難が発生する可能性がある

個人情報の盗難は、民事・刑事訴訟に発展する場合がある

企業は、従業員による公的あるいは未承認のAIモデルへのデータ漏洩防止のため、社内規定の整備を進めている

人間が作成したコンテンツは、オプトアウトしない限りトレーニングデータに使われる可能性がある

生成AIの使用に伴うリスクを特定する

誤った情報や有害な情報の拡散を防ぎ、自身や企業が金銭的・法的リスクにさらされないようにするために、人間による監視が不可欠である

自身が作成したコンテンツに対して、自らが責任を負うことを理解する

## 出題範囲 (参考訳)

- 有害なコンテンツや、民事・刑事訴訟の対象となり得るコンテンツを作成しない（いじめ、ヘイトクライム、詐欺、ストーカー行為、不正行為）
- 生成AIは、ディープフェイクなどの危険な目的に使用される可能性がある。本物のように見える有害・違法情報の生成が容易になる。個人情報の盗難など。

### 生成AIが社会に与える影響を特定する

#### マイナスの影響

- 人間同士の対話や交流が減少することによる影響を認識する
- 人間同士の関わりがAIに置き換えられる可能性を認識する
- AIへの過度な依存が人間のモチベーションに与える影響を認識する
- AIを利用して世論を操作しようとする人間の動機を認識する
- AIが私たちの仕事や人間らしさを奪うことへの恐れ
- 社会経済的な要因（誰もが平等にAIを利用できるわけではない）

#### プラスの影響

- 生成AIは、仕事の効率化に役立つ
- 特に異なる言語間での、より円滑なコミュニケーションを促進する
- より効果的な学習に役立つ
- 創造性を刺激するアイデアを生み出す。ブレインストーミングなど
- 日常的なタスクの効率化に役立つ（献立、レシピ、買い物リスト、友人や家族からの長文メッセージの要約など）
- パターンを分析し、機会として提示する
- 雇用を創出する（ただし従来とは異なる仕事）