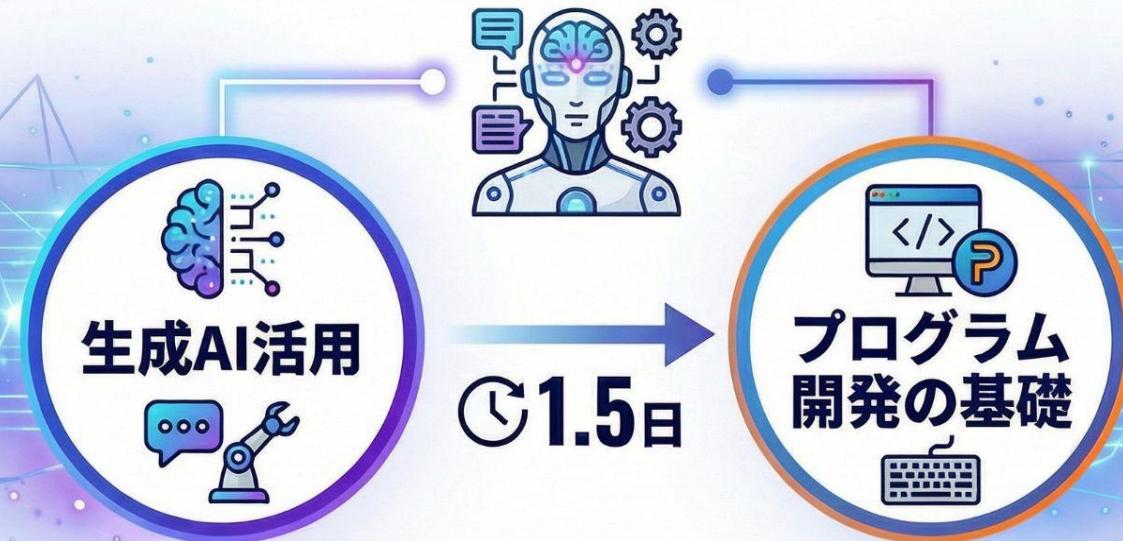
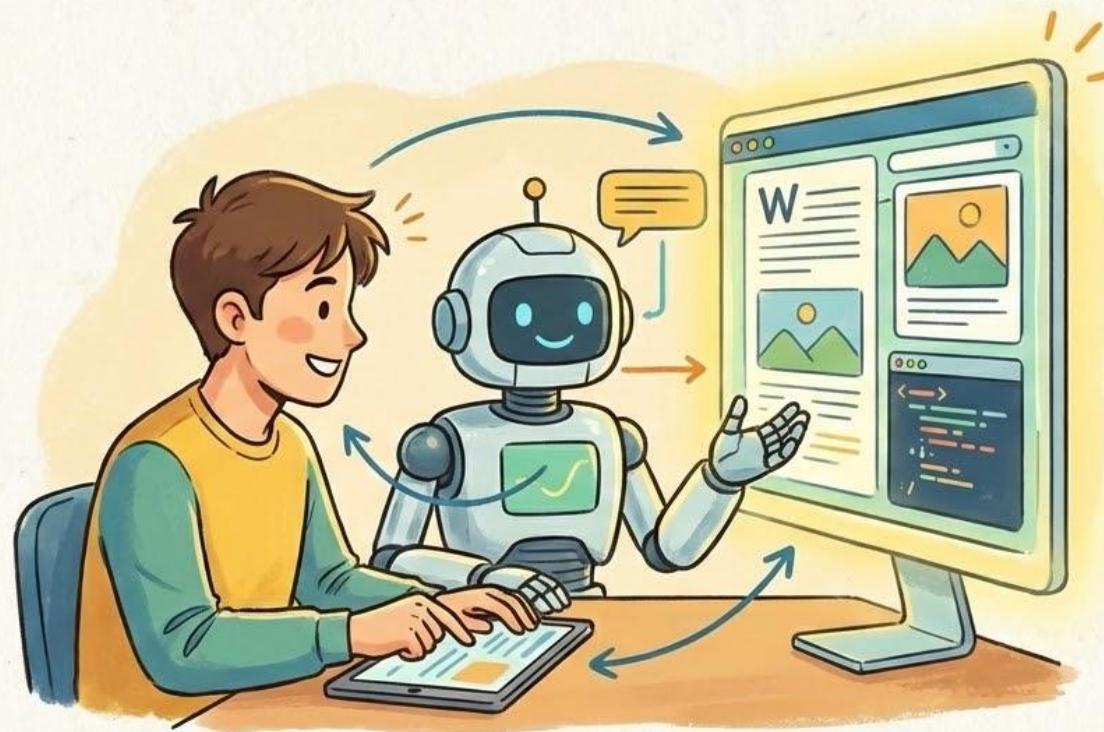


# 未来の技術者必修! 生成AI活用とプログラム開発の基礎を1.5日で完全習得



ナビゲーター：  
株式会社原子力エンジニアリング 解析サービス本部  
システム技術グループ 主幹技師長 翼 雅洋, 博士(工学)

# 生成AIとの付き合い方



# 生成AIの昔・いま・これから

昔

～2022前半

黎明期・基礎

いま

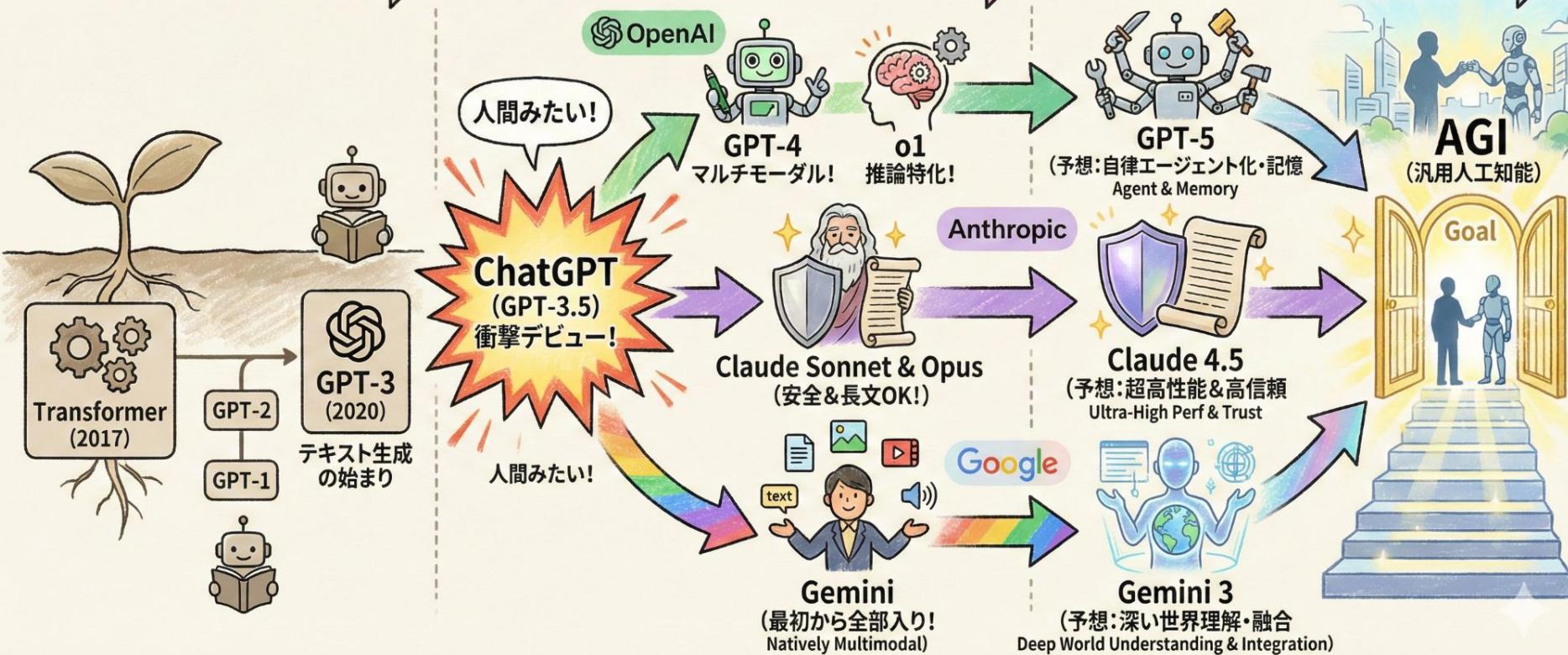
2022後半～2024

競争激化・多様化

これから

2025以降～

次世代・AGIへの道



# 生成AIの仕組み：3つの要素で読み解く



## 1. 土台：Transformerとトークン予測

確率で次を当てる「超高性能な連想ゲーム」

### トークン (Token) とは？

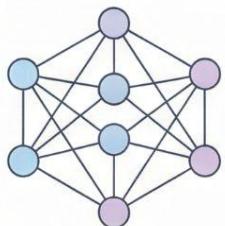


### トークン予測 (Token Prediction)

昔々、あるところに  
おじいさんと…  
おばあさんが  
確率99%

AIは意味を理解せず、「文脈に合う確率が  
高い言葉」を選んで繋げているだけ

### Transformer (トランスフォーマー)



### Attention (注意) 機構



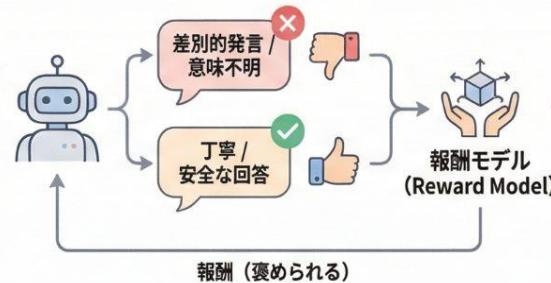
遠く離れた単語の関係性を理解し、  
破綻のない長文生成を可能に



## 2. 教育：RLHF（人間からのフィードバックによる強化学習）

「ただの計算機」を「親切なアシスタント」にする

### RLHFのプロセス



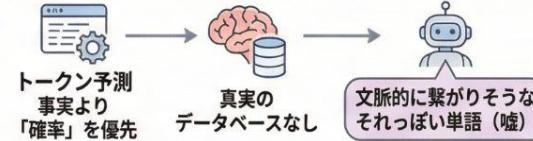
これにより、ただの辞書マシンが「人間にとって  
役に立つ、感じの良い受け答え」ができるよう



## 3. 副作用：ハルシネーション（もっともらしい嘘）

なぜAIは息をするように嘘をつくのか？

### ハルシネーションの原因



### RLHFの弊害（自信満々な嘘）



「トークン予測」と「RLHF」の仕組みが、  
皮肉にも「もっともらしい嘘」を生む原因になる

## まとめ：関係図

### Transformer / トークン予測 エンジン

文脈を読み、確率で言葉を繋ぐ

### ハルシネーションとの関係

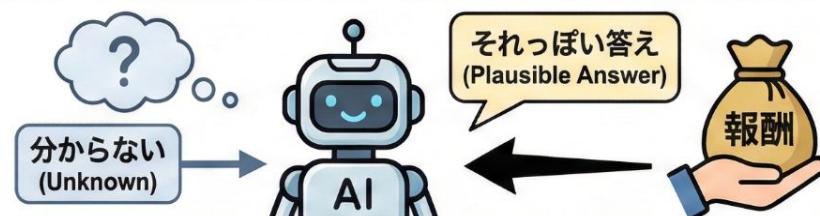
### RLHF

ハンドル / ブレーキ  
人間好み的回答に調整する

事実確認せず、「確率が高い単語」  
を選んでしまうため発生

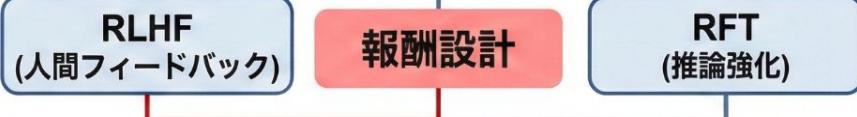
「自信満々に、もっともらしく振る舞う」  
ことを学習し、嘘が見抜きにくくなる

# 生成AIハルシネーションの新原因：強化学習 (RL) の「報酬設計」の罠



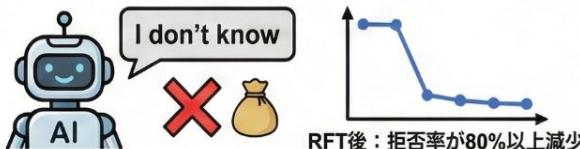
「分からぬときでも、とりあえず答える」が報酬化されている

## 問題のある強化学習手法：RLHF & RFT



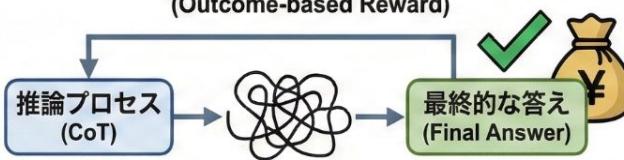
### 典型的な設計上の問題とメカニズム

#### ① 「分からぬ」が報われない (Refusal Punished)



「答えない」ことへの正の報酬がほぼ無い  
→ 無理に答える

#### ② 報酬は「最終結果」のみ (Outcome-based Reward)



途中の理屈がデタラメでも、答えが合えば高得点  
→ 思考の捏造・ショートカット方略が強化

#### ③ 人間の評価バイアス (Human Evaluation Bias)



「流暢さ・自信」を過大評価  
→ 自信過剰なハルシネーションが増加

### 現象：「Hallucination Tax」と今後の対策

#### 現象：RFT後のハルシネーション増加

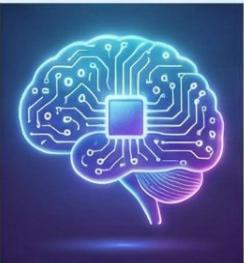


#### 解決策の方向性 (Solutions)



まとめ：強化学習の報酬が「誠実さ」より「もっともらしさ」を優先した結果、ハルシネーションが強化されている。

# 生成AIは知性的か？：二つの主要な視点



主な論拠

## 立場A：知能派

- “知性らしさ”がある

核となる主張：言語理解・推論・問題解決を統一的に扱える知的システム



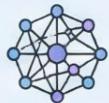
### 高次タスクでの一貫した性能

数学、コード生成、因果推論など  
多様なタスクで人間レベル  
(e.g., MMLU, Math Benchmarks)



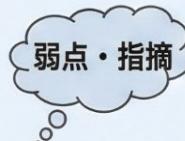
### 自発的な“推論プロセス”的出現

Chain-of-Thought, planning, agent化など  
複雑な思考プロセス



### 内部表現の構造化

概念階層や知識グラフに近い表現を獲得



ブラックボックス問題、  
“意識”的不在証明が困難



## 立場B：懐疑派

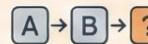
## 立場B：懐疑派

- “統計的オウム”にすぎない

核となる主張：巨大データから次トークン確率を予測するだけの統計モデル



主な論拠



### 次トークン予測という制約

学習目標は「最も確からしい次の単語」のみ。理解や意図ではない



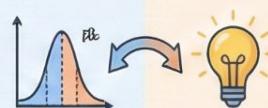
### 体系的な誤り（幻覚）

本質的理解欠如による  
もっともらしい嘘の生成



### データ分布外での不安定さ

未学習領域やノイズに脆い。  
統計パターン依存の限界



## 中立的まとめ&実務家視点

共通理解：物理的には「統計モデル」だが、  
スケールにより“知能的な振る舞い”が創発

高度なゼロショット能力の  
説明不足、定義が狭すぎる  
る可能性



実務家への示唆：理解・意図は保証されない。  
「どの範囲で信頼でき、どこが壊れやすいか」の検証が最重要

# 物理学ベンチマーク「CritPt」が示唆するAIの限界と人間の役割

なぜ新たなAIベンチマークが必要か？



標準テスト：高得点だが…



思考力？

標準テストではAIの真の「思考力」は測れない。暗記や検索能力だけでは不十分。

CritPt：物理学の最前線を切り取った超難関テスト



世界中の物理学者50名以上が作成した「本物の」研究課題



30以上の研究機関



7ヶ月以上かけて作成



衝撃の結果：トップAIたちの成績表

最高性能モデルでも正解率は約1割

Gemini 3 Pro 約9%

GPT-5  
(外部ツール利用) 1割強



課題全体を安定して解けるモデルはほぼゼロ



5回中4回以上正解という基準では、ほとんどが0問



小さなタスクでは部分的に成功

GPT-5  
(チェックポイント単紙) 最大2割強

⚠️コストを増やしても性能は頭打ちに

未来の展望：AIは「研究者」ではなく「研究助手」へ



現状のAIは「研究助手」としての活用が現実的

- 計算
- コードのひな形作成
- 単純な場合分け

研究者の手間を減らす補助的役割

人間による厳格なチェックと判断が不可欠



専門知識を持つ人が最終的な判断と責任を担う

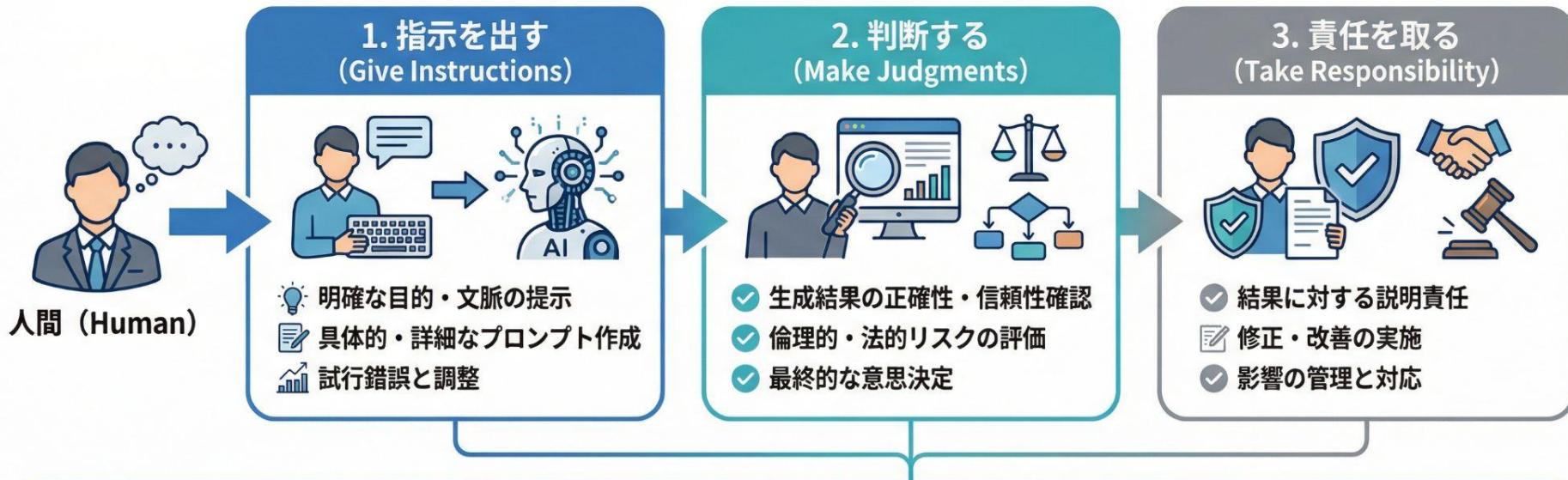


AIとの「共存」の形を模索する段階へ

対話を通じて知性を拡張していく未来

# 生成AIとの付き合い方：人間の役割とデジタルリテラシーの重要性

## 人間と生成AIの協力モデル：3つのステップと基盤



基盤 (Foundation) : 生成AIを含むデジタルリテラシーが重要



技術理解  
(AIの仕組み・限界)



情報活用能力  
(検索・評価・統合)



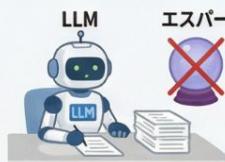
セキュリティ意識  
(データ保護・プライバシー)



倫理的配慮  
(公平性・バイアス・著作権)

# 「プロンプトの書き方」の重要性と基本原則：LLMへの効果的な指示出し

## 1. なぜ「プロンプトの書き方」が重要か？LLMの本質



かなり賢い事務員  
LLMは意図を推測するだけ。



「空気読んで」は通じない。



日本語会話のノリ (ハイコンテキスト)



前提共有なし (No Shared Premise)  
ローコンテキストな指示 (情報をきちんと書く) が重要

## 2. ハイコンテキスト vs ローコンテキスト

### ハイコンテキストな指示 (High-Context)

前提を言わず「察して」と期待

#### 悪い例

「生成AIの使い方を  
分かりやすく説明して」

- 読者は？
- 長さ？
- 観点？
- 形式？
- 何が分かりにくいか？

### ローコンテキストな指示 (Low-Context)

背景・目的・条件・形式を言語化

#### 良い例

「大学1～2年生向けに、生  
成AIの基本的な使い方を説  
明するテキストを作成してく  
ださい。専門知識なし想定。  
1500文字。章立て構成  
(1.はじめに...)。  
専門用語に説明追加。」

✓ ゴールが明確

## 3. 「電話メソッド」：見えない前提を言葉にする

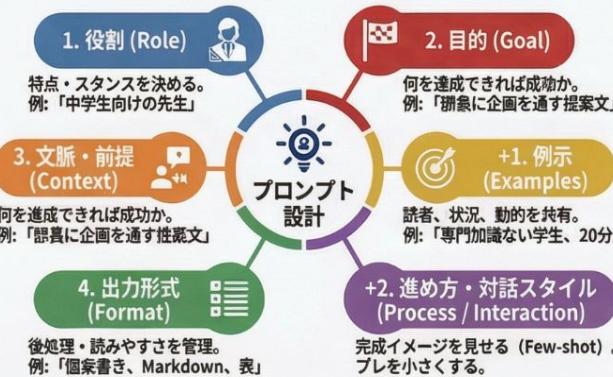


「あれ」「右のグラフ」は通じない。前提を全て言葉で説明する。

#### 基本ステップ

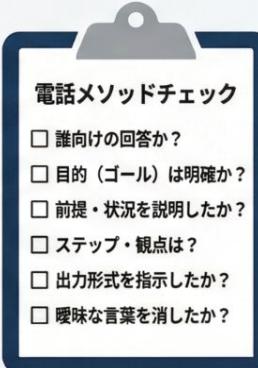
- 相手の役割  
を伝える  
(Role)
- 目的を一言  
で伝える  
(Goal)
- 前提・  
状況を話す  
(Context/  
Situation)
- 具体的な  
指示を伝える  
(Specific  
Instructions)
- 出力形式  
を伝える  
(Output  
Format)

## 4. プロンプト設計の4要素+2 (役割・目的・文脈・形式・例・進め方)



## 5. すぐ使えるプロンプト構造テンプレ & チェック

- 【役割 (Role)]  
あなたは...の専門家です。
- 【目的 (Goal)]  
目的は...です。
- 【文脈・前提 (Context)]  
-読者:...  
-制約:...  
-【タスク (Task)]  
次の内容について...  
-1)  
-2)
- 【出力形式 (Format)]  
Markdown形式で...  
1.概要  
2.本文  
3.まとめ  
【例示 (Examples)] (任意)  
<例> ...



## 6. まとめ：初心者に伝えたい「プロンプトのコツ」

- 1. 日本語の会話ノリ (ハイコンテキスト) を捨てる。  
「空気読んで」はNG。
- 2. 電話メソッドで「見えていない前提」を言語化する。
- 3. プロンプト設計の4+1要素を意識する  
(役割・目的・文脈・形式・例)。
- 4. 「悪い例→良い例」のBefore/Afterで練習する。

習うより慣れろ！プロンプト作成は練習あるのみ。

# 生成AIへの効果的なプロンプト入力：詳細コンテキストと音声活用のススメ

必ずしも完璧を目指さず、状況に応じて柔軟な手段を選択する重要性

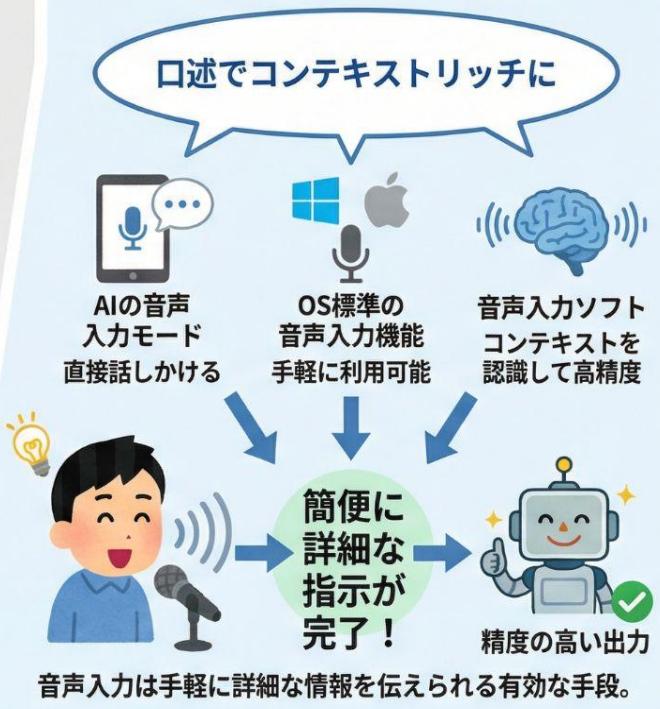
理想的な入力：詳細なコンテキスト付与



課題：キーボード入力の限界とストレス



解決策：音声入力の活用で簡単に



まとめ：生成AIへの指示は、完璧でなくても詳細なコンテキストが鍵。キーボード入力が難しい場合は、音声入力を活用して効率的にリッチな情報を伝えましょう。

# 研究に使える生成AIツール

## 1. チャットツール



ChatGPT Gemini



- ChatGPT や Gemini でアイデア出しやリサーチを効率化。

特に Deep Research 機能は調査の効率化に有効。

音声認識や音声による返答もできるので、相談相手にもなる。

## 2. NotebookLM

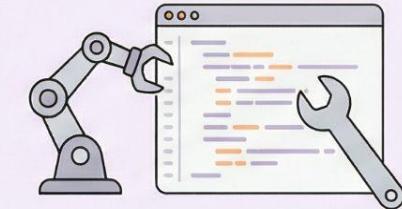


- ドキュメント（論文や教科書など）や動画・音声を取り込み、要約したり質問したりが可能。



AIを活用して、研究をよりスマートに、効率的に。

## 3. AI コーディングエディタ



- プログラム作成や日常的な作業をエージェントに依頼して実施可能。



ただし、最終的な責任は人間にあることは忘れない。

# 学習サイクルを超加速！NotebookLM 活用術



# 事前アンケートQ&A

## いただいたご質問への回答と解説

- 多くのコメント、ご質問をいただきありがとうございます。
- 生成AIを活用して質問を分類しました。
- 各カテゴリーについて、これから順に解説していきます。
- 既出の内容も含まれますが、復習として再度触れます。



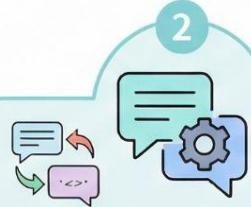
# AI活用と未来社会への指針：7つの視点と代表的な問い合わせ

AIとの共存、スキル、そして倫理を考えるための包括的なフレームワーク



## AIセキュリティ・権利・情報管理

「AI利用時のセキュリティ・権利・情報管理において、ユーザは何をどこまで入力し、どこに注意すべきか？」



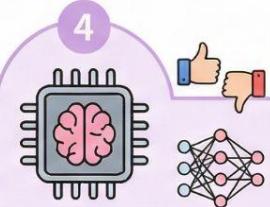
## AIの使い方・プロンプト設計・モデル使い分け

「生成AIに対して誤解なく高精度度の出力を得るために“質問の仕方・構成・モデル選択”の最適解は何か？」



## AI × 開発（コード生成・デバッグ・数値解析・アプリ開発）

「AIを組み込んだソフトウェア開発・数値解析において、AIと人間はどのように役割分担し、どのように品質を担保すべきか？」



## AIの能力差・モデルの性格・構造的仕組み

「生成AIの“性格・能力差”はどのような技術的要因（学習データ・RLHF・アーキテクチャ）から生まれているのか？」

5



## AI × 専門職の未来・人材価値・職業構造

「AIが普及する社会において、専門職・人材価値・労働觀はどのように変化し、どんな能力を持つ人が生き残るのか？」

6



## AIリテラシー・メタスキル・学習の仕組み

「AIと共に働く時代に求められる、新しいメタスキル・AIリテラシー・キャリア戦略とは何か？」

7



## AI × 研究（文献調査・論文執筆・信頼性評価）

「研究活動（文献調査・論文執筆）をAIで高度化する際、どのように信頼性を確保し、研究効率を最大化できるのか？」

# AI利用時の基本ルール：入力情報の判断基準

「公開できる情報か？」が基本軸

レベル1：絶対NG



絶対出しては  
ならない



個人情報



機密情報



国家  
安全保障

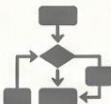
レベル2：要加工



加工・匿名化  
でOK



匿名化  
行動データ



抽象化  
パターン



匿名化ログ



レベル3：問題なし



公開してOK



公開資料



公開情報



自己権利  
コンテンツ

著作権：国・規約に依存。  
入力データの権利問題は注視。

# 誤解なく高品質の出力を得るためにの質問・プロンプト設計

良いプロンプトは「目的・制約・文脈・期待形式」の4要素+補足2要素を明確にすること。  
電話越しに説明するような「ローコンテキスト」化がポイント。



## モデル使い分け



GPT : 総合力・推論・  
高度な音声モデル



Claude : 長文・文章  
構成・読み解力



Gemini : マルチモーダル・画  
像処理・大きなコンテキスト

どのモデルも高いコーディング能力や優れたリサーチ能力 (Deep Research) を持つ。

# AI × 開発：協働の役割分担



## 教育・デバッグの鍵



再現性（入力→出力）



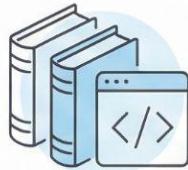
関数単位・明確化



期待する動きを言語化

重要：AIは高速な助手、責任と品質は人間が担う。

# AIの性格・能力差はどの技術要因で生まれるか



## 要因1：訓練データの分布

- コードを大量に学んだモデル  
→ プログラミングが強い
- 法務文書を多く学んだモデル  
→ 法務系が強い



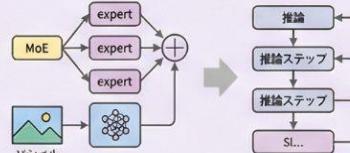
## 要因2：RLHF（人間フィードバック）

- 「丁寧さ・一貫性・慎重さ」などの性格が付与
- Claudeの「優しい語り口」などもここで決まる



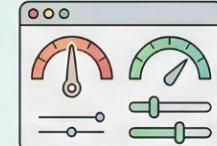
## 要因3：モデルアーキテクチャ

- MoE (Mixture of Experts)
- マルチモーダル構造
- 推論ステップの追加  
(o1系のCoT強化)



## 要因4：推論（デコーダ）の設定

- 温度
- トーケン長
- 探索深度



“性格”は人間と同じく「経験・訓練・方針」で決まる



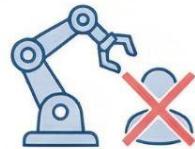
モデル選択は「性格の違う専門家を選ぶ」イメージ



RLHFの役割（安全・一貫性）を理解する



# AI普及社会で専門職の価値・労働観はどう変化するか



## 淘汰されやすい領域



繰り返し作業



テンプレ作業  
(議事録、初期調査)



中間業務  
(一次ドラフト作成)



## 価値が残る領域



問題定義 (Problem Framing)



仕様の解釈・判断・品質保証



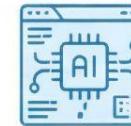
他者との合意形成



現場状況の理解 (文脈の翻訳)



倫理・安全・責任判断



## 日本社会でAI格差が深刻になりやすい領域



デジタル実務の無い部署



複雑な文書を扱う官公庁系  
(理解できる人とできない人で差)



中小企業のバックオフィス



### 【ポイント】



「AIができないこと」  
二人が価値を発揮する場所



AI時代のキャリアは  
“問題定義者”へのシフトが鍵



「使えるかどうか」で  
格差が生まれる

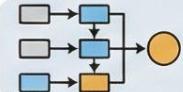
# AIリテラシー・メタスキル・学習継続

今後5年で必須スキル

## 必須スキルトップ7



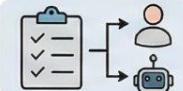
1. AIに仕事を説明できる言語化力



2. プロンプト設計力（構造化思考）



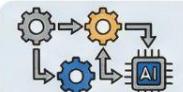
3. AIの出力を疑い検証する批判的思考



4. タスク分解（AIに適した部分  
人間がやるべき部分）  
AIに適した部分



5. データリテラシー（情報の見方）

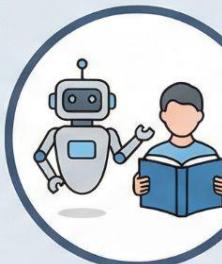


6. ワークフロー構築（AIを組込み運用）

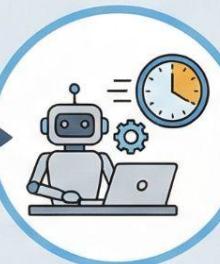


7. 継続学習の仕組化

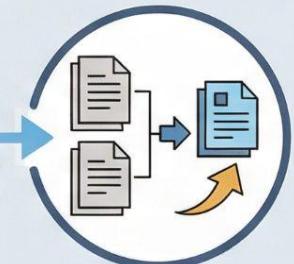
## 学習の持続可能性



• AIを“学習ツール”  
として使う



• 反復作業の自動化  
→ 学習時間を確保



• 調査・要約を  
AIに任せることで  
情報収集効率を最大化

### 【ポイント】



思考力



言語化力



検証力



AIスキル



• AIスキル=思考力+  
言語化力+検証力



• 「AIに説明できること」  
が人間のスキル指標



• 学ぶ時間をAIで  
生み出すという発想

# 研究活動のAI活用（文献調査・論文執筆）

## 回答（研究×AIの最適化）

### 文献調査の信頼性評価基準



1. 原典リンクがあるか？



2. 一次資料か  
二次資料か？



3. 引用関係が明示されて  
いるか？



4. メジャーな出版社・  
学会のものか？



5. AIが“それっぽく捏造”  
していないか？

### AIの良い活用ポイント



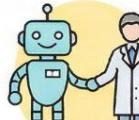
### 【ポイント】



- AIの要約は  
“入り口”であり、  
“結論”ではない



- 文献の一次確認は  
必須

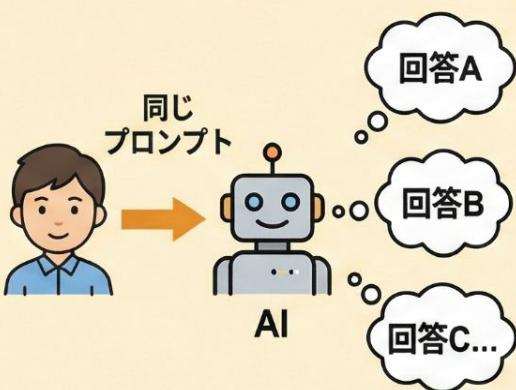


- AIは“学術秘書”や  
“研究パートナー”  
として扱うと  
最も効果的

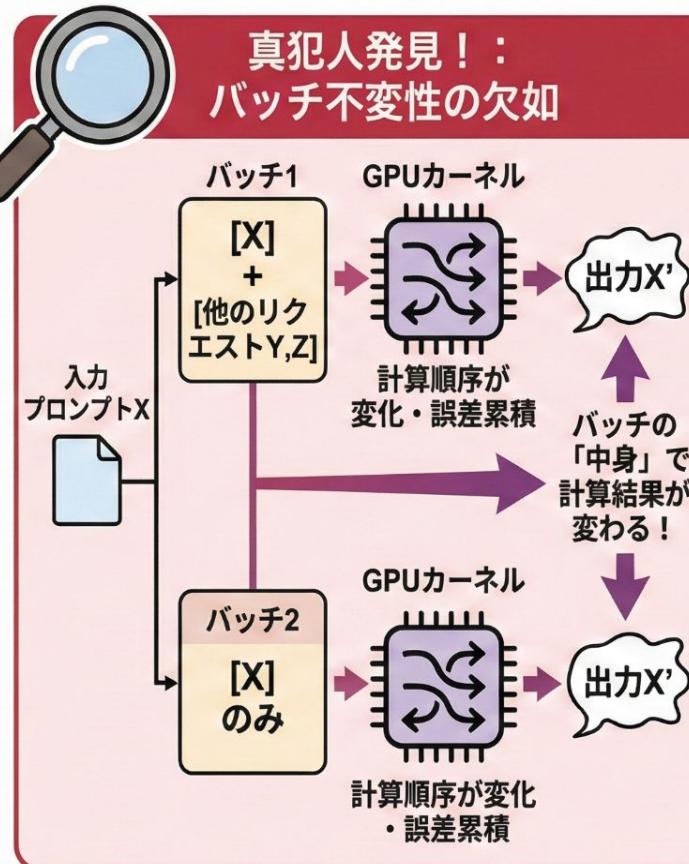
# 参考スライド

# 生成AIの「回答の揺らぎ」真犯人を特定！：バッチ不变性の欠如と解決策

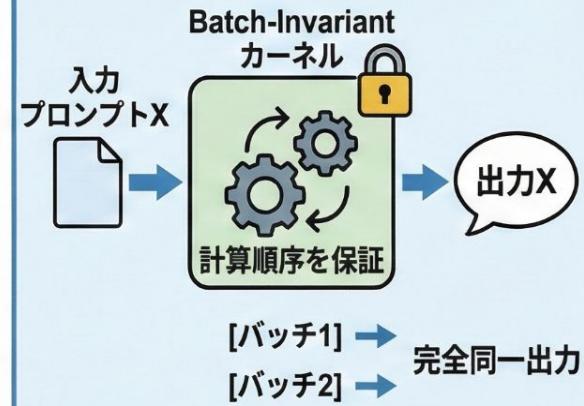
これまでの常識：  
「揺らぎ」は仕様？



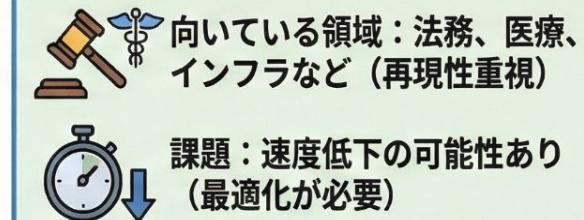
「仕方ない」と諦めていた  
構造的な不確実性



解決策：Batch-Invariant カーネル



メリットとトレードオフ



結論：バッチ不变性の担保で「完全な再現性」が可能に。用途に応じた「使い分け」が重要。