

第1章 AIの現在地

～2025年 知っておくべきメガトレンドと基本知識～

学ぶこと：AIの主要トレンド

AIの現状理解のための主要技術トレンドと基本概念。

- **AIの定義と分類**
- **自然言語処理技術 (Transformer, LLM)**
- **データと機械学習**
- **AIアラインメント**
- **生成AI**
- **AI for Science**
- **AIの効率性と持続可能性**

現代社会におけるAIの重要性

AIは社会の様々な場面で活用。

DCON等では、学生のアイデア実現の強力な手段。

本章でAI活用の基礎知識を習得。

1.1 AIの定義と分類

AI（人工知能）とは

人間の知的活動をコンピュータで模倣・実現する技術・学問。

AIの主な分類

- **特化型AI (ANI)** : 特定タスク特化。現在の主流。
例：顔認証、自動翻訳。
- **汎用AI (AGI)** : 人間のように多様なタスク対応。未実現。

ポイント : 現在のAIは「特化型」。万能ではないが特定分野で高性能。

1.2 Transformerと大規模言語モデル(LLM)

自然言語処理のブレークスルー

Transformer (2017年): NLPに革命。 **自己注意機構** で深い文脈理解。

大規模言語モデル (LLM)

Transformerベース。膨大なテキストデータで事前学習。例: GPTシリーズ。

LLMの主な能力

高度な文脈理解、文章生成、翻訳、対話、要約。

キーワード: Transformer, LLM, 自己注意機構, GPT, NLP

LLMの限界と課題

- **ハルシネーション** : 誤情報生成。
- **バイアス** : 学習データの偏見を反映・増幅。
- **知識のカットオフ** : 最新情報に未対応。
- **計算資源とエネルギー消費** : 膨大なリソース。

批判的視点での活用が重要。

1.3 データ駆動型AIと機械学習

AIとデータ

機械学習モデルはデータから学習。データがAI性能を左右。

機械学習の主要な手法

1. **教師あり学習** : 入力と正解ラベルで学習。例：迷惑メール分類。
2. **教師なし学習** : 正解ラベルなしでデータ構造発見。例：顧客セグメンテーション。
3. **強化学習** : 試行錯誤で報酬最大化。例：ゲームAI。

キーワード: 機械学習 (ML), 教師あり学習, 教師なし学習, 強化学習, 訓練データ

学習データの重要性：質・量・バイアス

- **データの量**：多いほど汎化性能向上。大規模モデルでは特に重要。
- **データの質**：不正確・ノイズは性能低下 (GIGO)。**データバイアス** は判断に偏り。

高品質で偏りのないデータ準備が不可欠。

1.4 AIアライメント：人間との価値整合

AIアライメントとは

AIの行動・目標を人間の価値観・意図・社会全体の利益と一致させること。

- **指示チューニング**：指示と望ましい出力ペアで学習。意図理解向上。
- **RLHF**：人間評価を報酬にLLM応答を人間の好みに近づける。

AI高度化に伴い重要性増。

キーワード: アライメント, RLHF, 指示チューニング

1.5 生成AI：新たなコンテンツ創造

生成AI (Generative AI) :

既存データから学習し、新規オリジナルコンテンツを自動生成。

- **テキスト生成** : LLM (GPT等)、プログラムコードなども
- **画像生成** : 拡散モデル (Stable Diffusion等)
- **音楽生成**
- **動画生成** (Sora等)
- **3Dモデル生成**

マルチモーダルAI : 複数情報（テキスト、画像等）を統合処理・生成 (GPT-4V等)。

| 生成AIの活用例 (DCONを想定)

- アイデア創出支援
- 資料作成補助
- デザインプロトタイピング
- コーディング支援

生成AI活用で創造活動を加速。

1.6 AI for Science：科学研究の革新

AI for Scienceとは

AIを科学研究に応用。新発見加速、課題解決。大量データ解析、シミュレーション高速化、仮説生成。

主な応用分野の例

- **生命科学・創薬**：AlphaFold
- **材料科学**：マテリアルズ・インフォマティクス
- **地球科学**
- **天文学**

1.7 AIの効率化と持続可能性

AIの環境負荷

大規模AIモデルは訓練・運用に多大な電力消費。環境負荷が課題。
グリーンAI、サステナブルAIが求められる。

主な取り組み

- モデル軽量化・圧縮
- 効率的なAIアーキテクチャ
- 省エネAIハードウェア
- エッジAI

1.8 責任あるAI (Responsible AI)

責任あるAIとは

AIが人間の価値観・社会規範に沿い、公平・透明・安全・説明可能に運用されることを目指す概念。

主要原則

公平性 , 透明性 , 説明可能性 (XAI) , 安全性と信頼性 , プライバシー保護 , アカウンタビリティ , 人間中心性

キーワード: 責任あるAI, ELSI, 公平性, 透明性, XAI, AI倫理

まとめ：AIの現在地と今後の展望

本章のまとめ：

- AI定義・分類、**Transformer** と **LLM** の影響。
- **データ** の重要性（質・量・バイアス）。
- **AIアラインメント** の必要性。
- **生成AI**、**AI for Science** の進展。
- AIの **効率化・持続可能性**、**責任あるAI**。

これらはAI理解・活用の基礎。AIは社会を変革し続ける技術。

質疑応答

ご質問、さらに詳しく知りたい点など、ご遠慮なくどうぞ。