

生成AIパスポート試験：第2章 要点まとめ

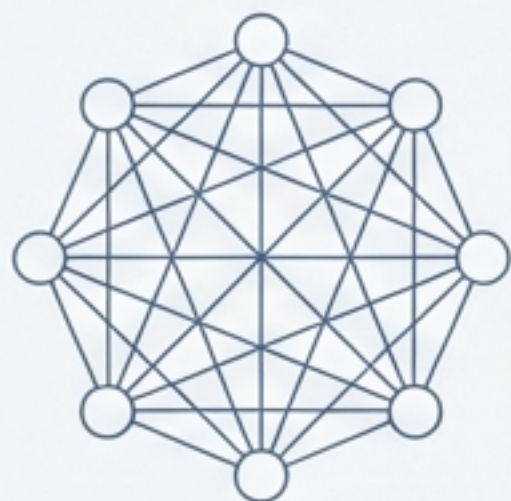
2026年シラバス対応：ボルツマンマシンから自律型エージェントまで



生成モデルの起源：確率的モデルの進化

Boltzmann Machine

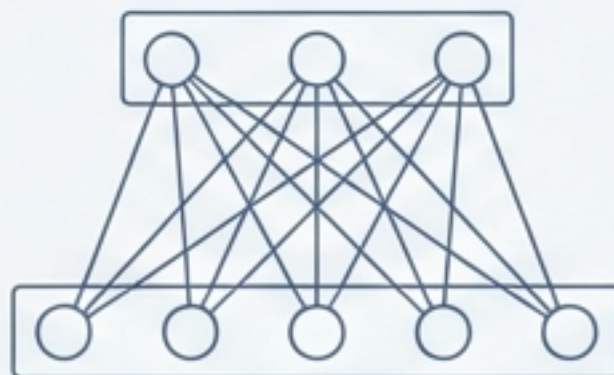
1980s | Hinton & Sejnowski



確率的に「最良の答え」を探索するニューラルネットワーク。
弱点：データ量が増えると処理時間が爆発的に増加。

Restricted Boltzmann Machine (RBM)

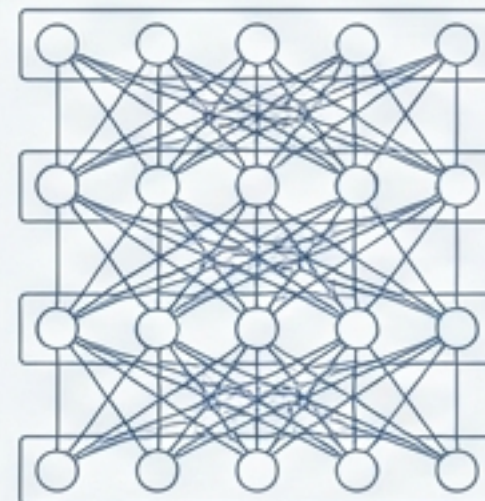
1986 | 効率化の実現



ニューロン間の接続を制限し、
入力層と推定層を分離。
革新：教師なし学習が可能に。
計算コストを大幅削減。

Deep Learning

2000s | 深層学習



RBMを多段的に重ねた「深層信頼区間ネットワーク」。
人間の脳を模倣し、複雑なパターン認識を実現。

時系列データの克服：RNNからLSTMへ

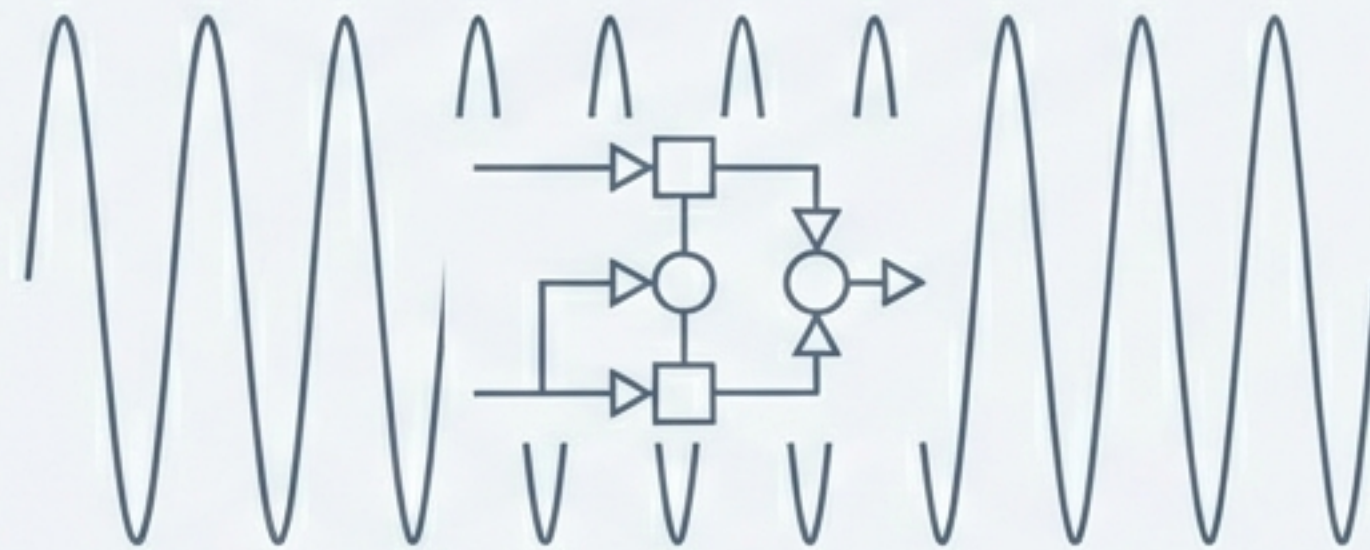
自己回帰モデル (Autoregressive Models) の進化

RNN (再帰型ニューラルネットワーク)

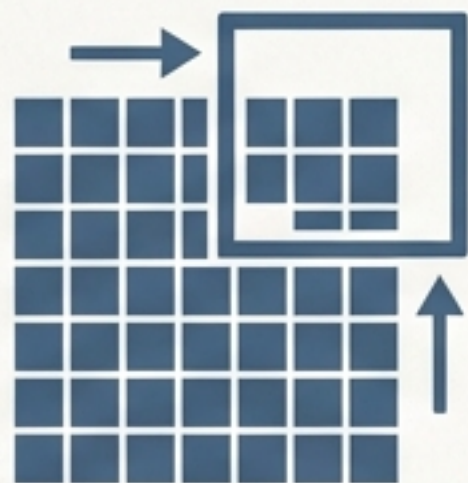


時系列データ (株価、天気、音楽) の予測に使用。
弱点：過去の情報を保持できず、長期的な文脈を忘れてしまう。

LSTM (長短期記憶)



画像と構造の理解：Transformer以前の主要モデル



CNN

畳み込みニューラルネットワーク

役割：画像認識・処理

仕組み：畳み込み

(Convolution) により、局所的な特徴から全体像を抽出。

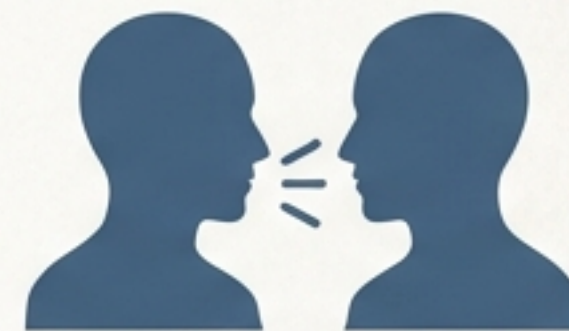


VAE

変分オートエンコーダ

役割：圧縮と復元

仕組み：Encoder (圧縮) → 潜在変数 → Decoder (復元)。入力に「似た」データを生成。



GAN

敵対的生成ネットワーク

役割：高品質な生成

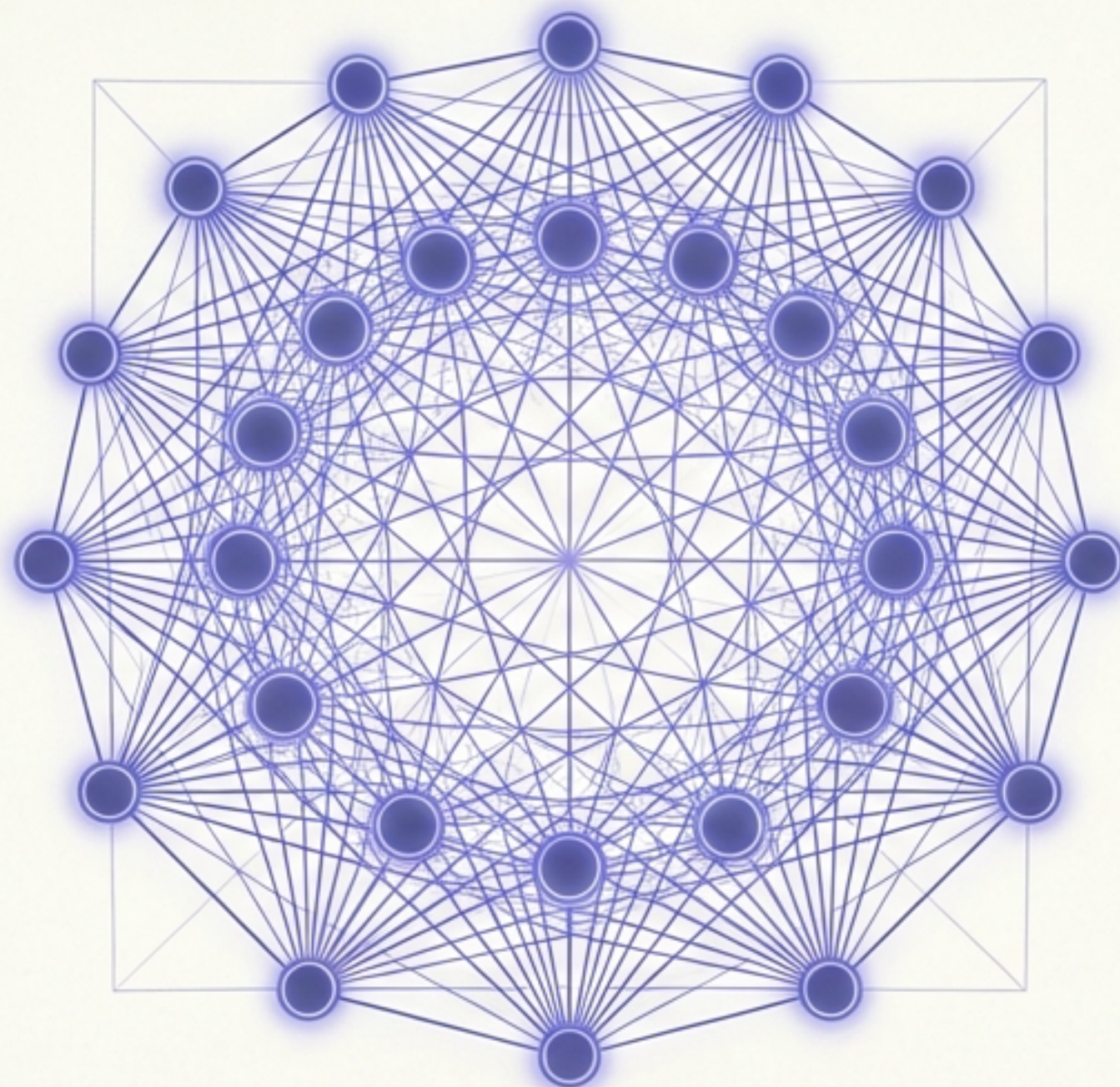
仕組み：生成器 (偽造者) vs 識別器 (警察)。競い合うことで精度を極限まで高める。

2017年の革命：トランスフォーマーモデル

自己注意機構（Self-Attention）によるパラダイムシフト

革新技術：Self-Attention （自己注意機構）

- 定義：文中のすべての単語の重要度を、順序に関係なく一度に計算する仕組み。



Self-Attention Web

パラダイムシフト

BEFORE (RNN):
単語を1つずつ順番に
処理 → 遅い

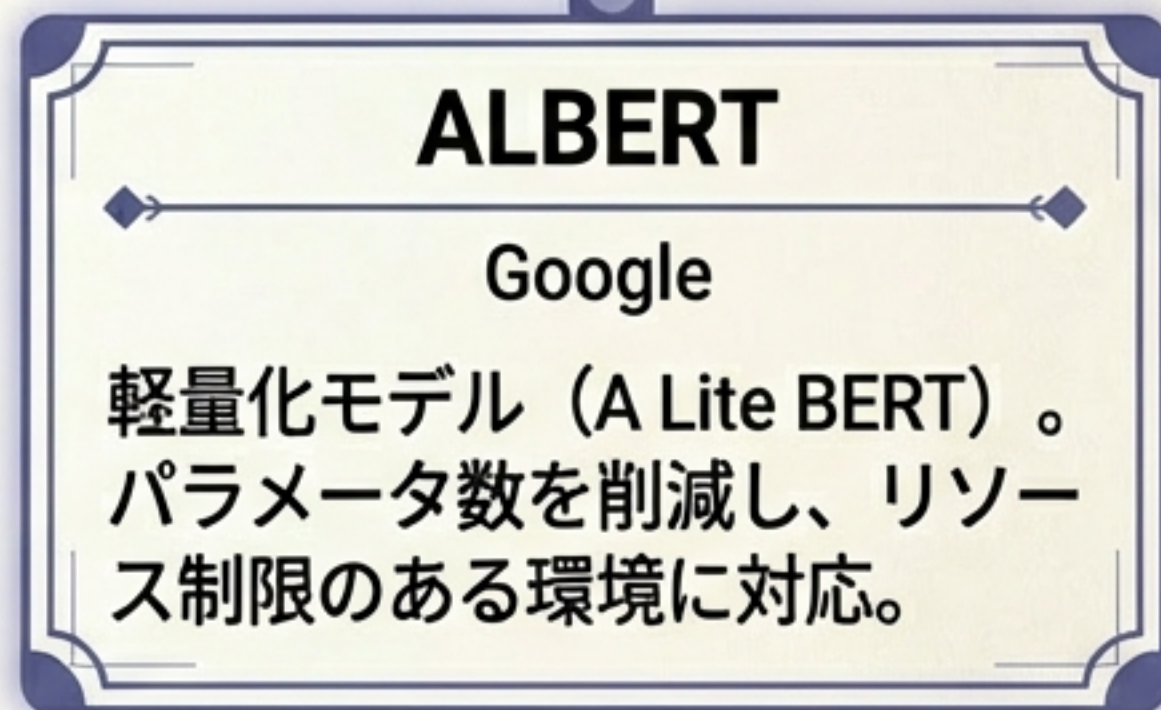


AFTER (Transformer):
全単語を並列処理
(Parallel Processing)
→ 劇的に高速

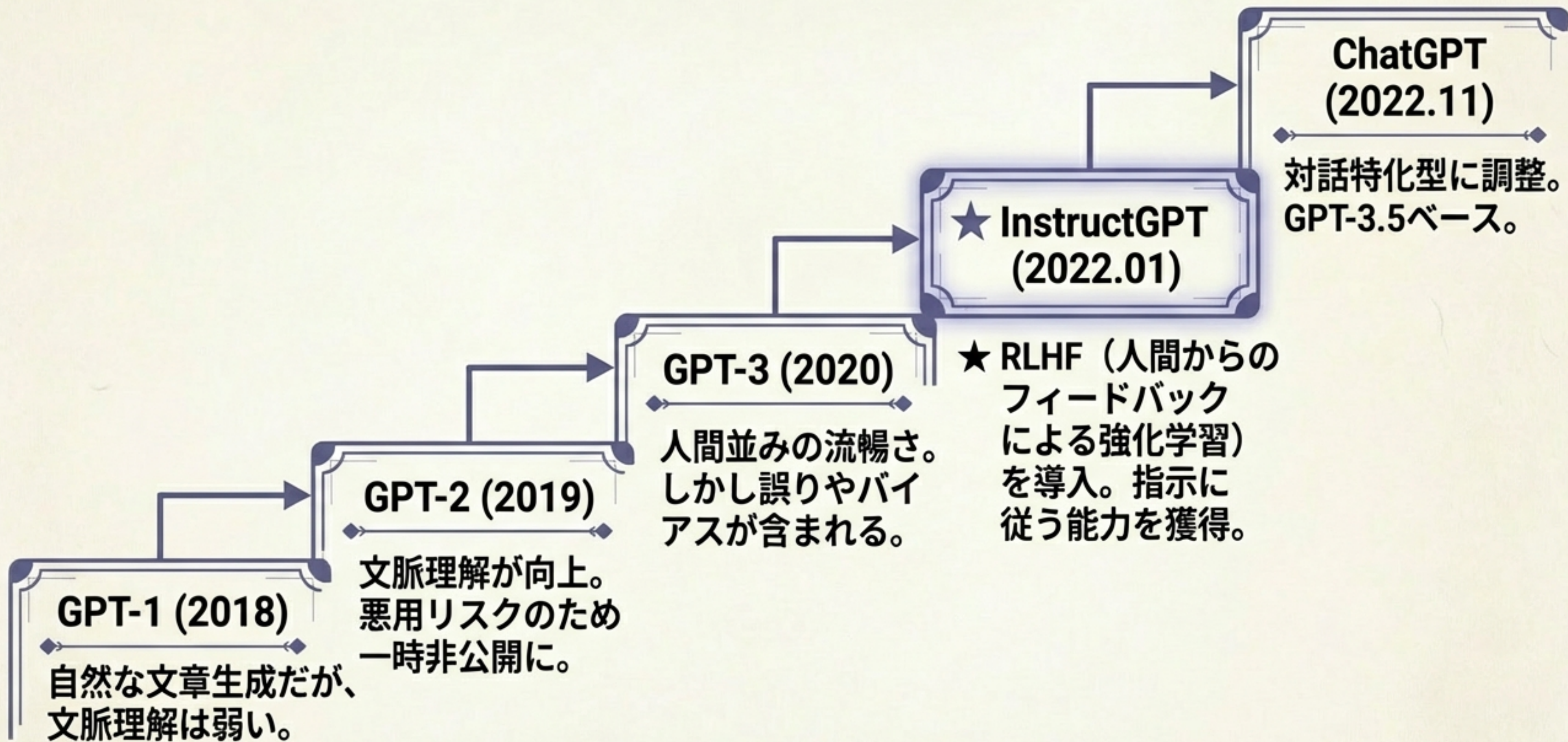
RNN/CNNへの依存を脱却し、現代の大規模言語モデル（LLM）の基礎となる。

文脈理解の深化：BERTの系譜

双方向 (Bidirectional) エンコーダモデルの派生



GPTの進化：テキスト生成から対話へ



マルチモーダルとエコシステム：GPT-4

GPT-4 Architecture

- マルチモーダル：テキスト、画像、音声、動画を同時に処理
- ハルシネーション（嘘）の低減
- 24言語以上に対応
- GPT-4o：リアルタイム音声対話を実現（Omni-model）

The Ecosystem

Code Interpreter

Pythonコードを生成・実行。データ分析やグラフ作成が可能。

GPTs

特定の目的にカスタマイズされたChatGPTを作成・公開。

推論重視型モデル：oシリーズ

「考える時間」を持つAI

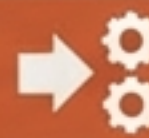
Standard GPT

Fast Response

o-Series
(Reasoning)



Thinking Process (Chain of Thought)



High Accuracy Response

GPT-o1

論理的なタスクで安定した精度を発揮。

GPT-o3 (Future Syllabus)

さらに深い思考を行うモデル。

- バリエーション：
 - o3-mini (標準)
 - o3-Pro (高コスト・最高精度)

次世代フラッグシップ：2025-2026 (シラバス準拠)



GPT-4 (Deliberative)

自律的なツール使用（検索、Python）が可能。

GPT-4.1

開発者向け強化版。コーディング指示の遵守とバグ修正能力の向上。

GPT-5

- 推論速度の大幅向上
- ハルシネーションの劇的な減少
- Pro版の提供開始

自律型エージェントと専門ツール

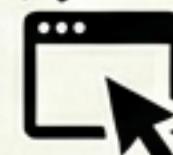
Sora (Video)

2024.02 | テキストから動画生成
最大60秒、1080pの高画質映像。



Operator (Web Agent)

2025.01 | ブラウザ操作代行
検索、予約、購入などのWebタスク
を計画・実行。



Codex (Software Agent)

2025.05 | 開発エンジニアリング
バグ修正から実装、テストまで一貫し
て行う。

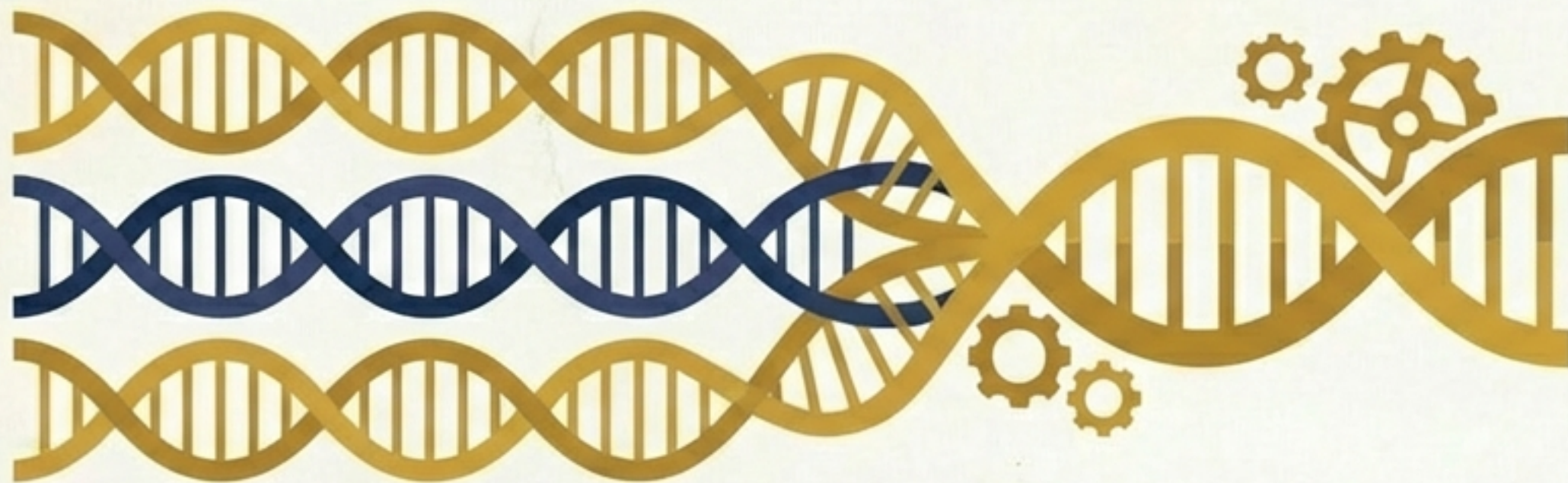


Image Gen (Autoregressive)

左上から右下へ描画する自己回帰方式。
画像内の文字描写に強い。



競合分析：Google Gemini エコシステム



Native Multimodal

最初からマルチモーダルとして設計（継ぎ接ぎではない）。

Gemini Nano

オンデバイス（スマホ端末内）で動作する軽量モデル。

Gemini Pro / Flash

バランス重視の標準モデル。

Gemini Ultra / 2.5

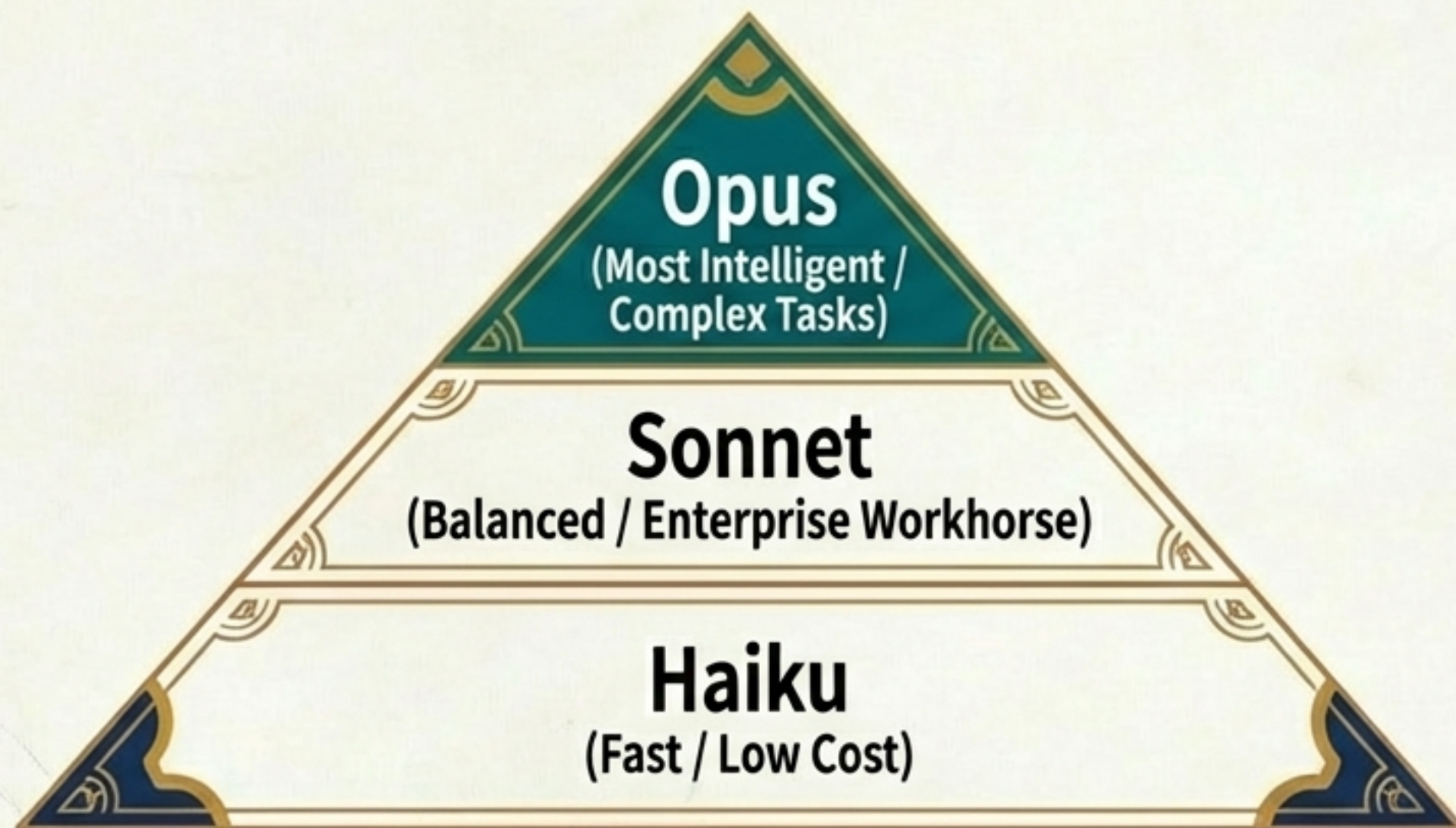
推論量を自動調整する最高性能モデル。

1.0 → 1.5 → 2.5へと進化

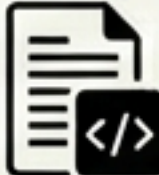

競合分析：Anthropic Claude

Philosophy: Constitutional AI (コンスティテューショナルAI)

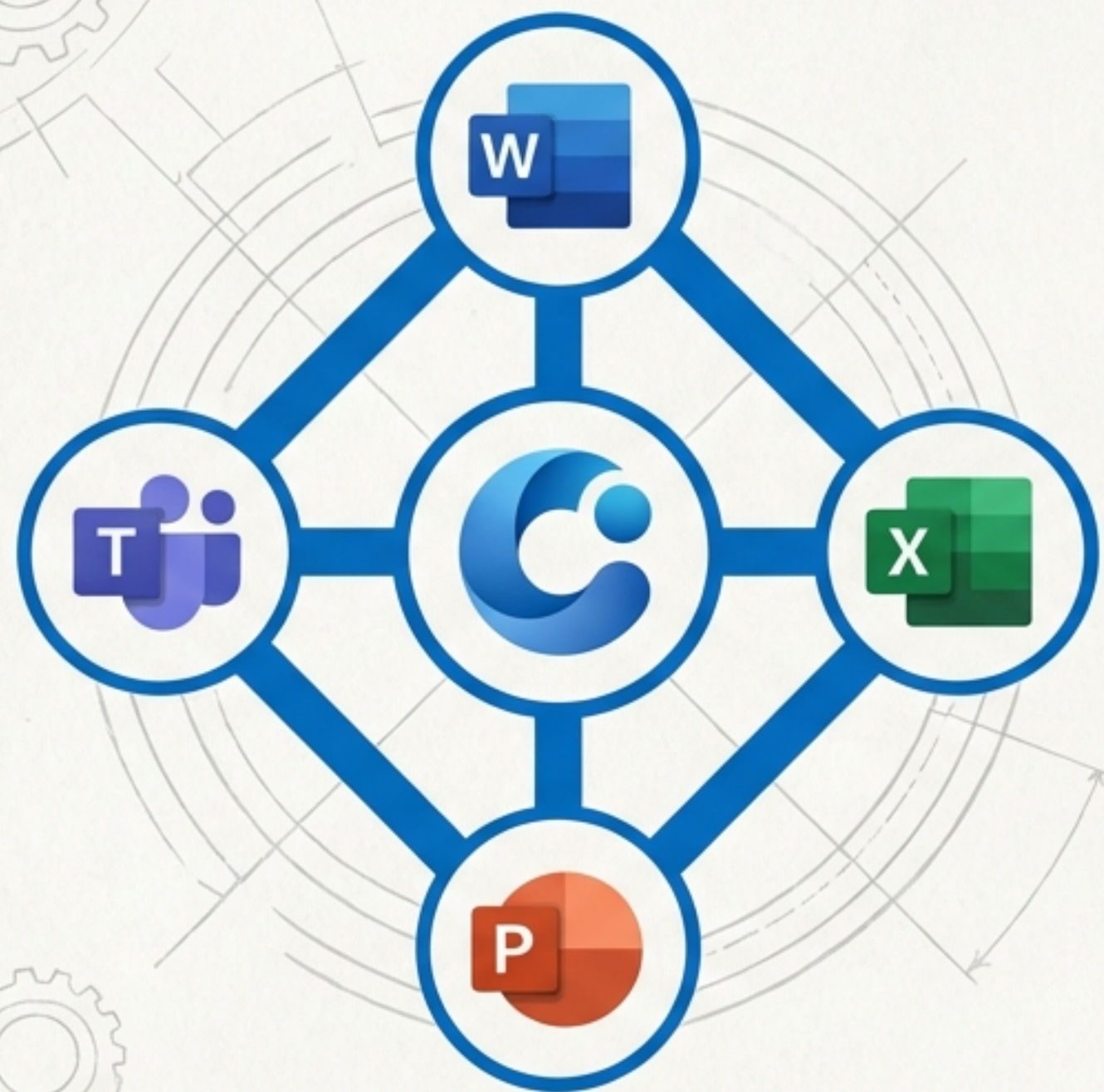
無害・役立つ・正直 (Harmless, Helpful, Honest) を原則とする。



New Features

-  **Artifacts:** コードやドキュメントを別ウィンドウで表示・編集
-  **Claude 4 (2025):** 外部ツールを並列使用するハイブリッド設計

競合分析：Microsoft Copilot



Engine: OpenAI GPT-5 (Based)

Copilot PC

機能「Recall」：過去の画面操作や作業履歴をオフラインで検索。

GitHub Copilot

プログラミングコード生成支援。

Security Copilot

セキュリティ分析と対策のサポート。

試験直前チェックリスト：重要用語

- ✓ **RLHF (Reinforcement Learning from Human Feedback)**
人間からのフィードバックを用いた強化学習。ChatGPTの自然な対話の鍵。
- ✓ **Multimodal (マルチモーダル)**
テキスト、画像、音声を一度に処理する能力。
- ✓ **Hallucination (ハルシネーション)**
もっともらしいが事実ではない嘘の情報を生成する現象。
- ✓ **Self-Attention (自己注意機構)**
並列処理を可能にし、学習時間を短縮させたTransformerの核。
- ✓ **PEFT (Parameter Efficient Fine-Tuning)**
パラメータ効率の良いファインチューニング (LoRA/ALBERT等の概念)。

進化の系譜を暗記せよ： 統計的モデル → Deep Learning → Transformer → Agent