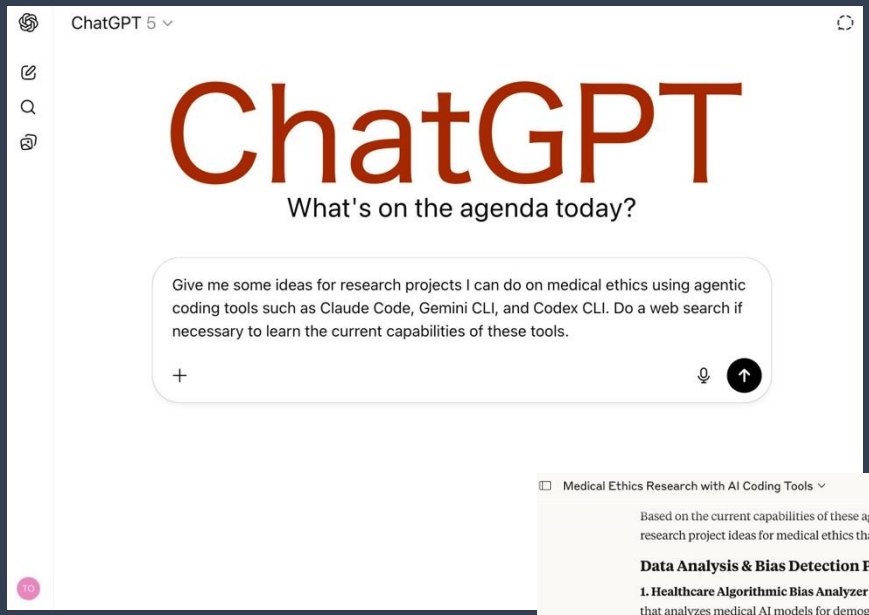
AIエージェント時代における研究の卓越性の再定義

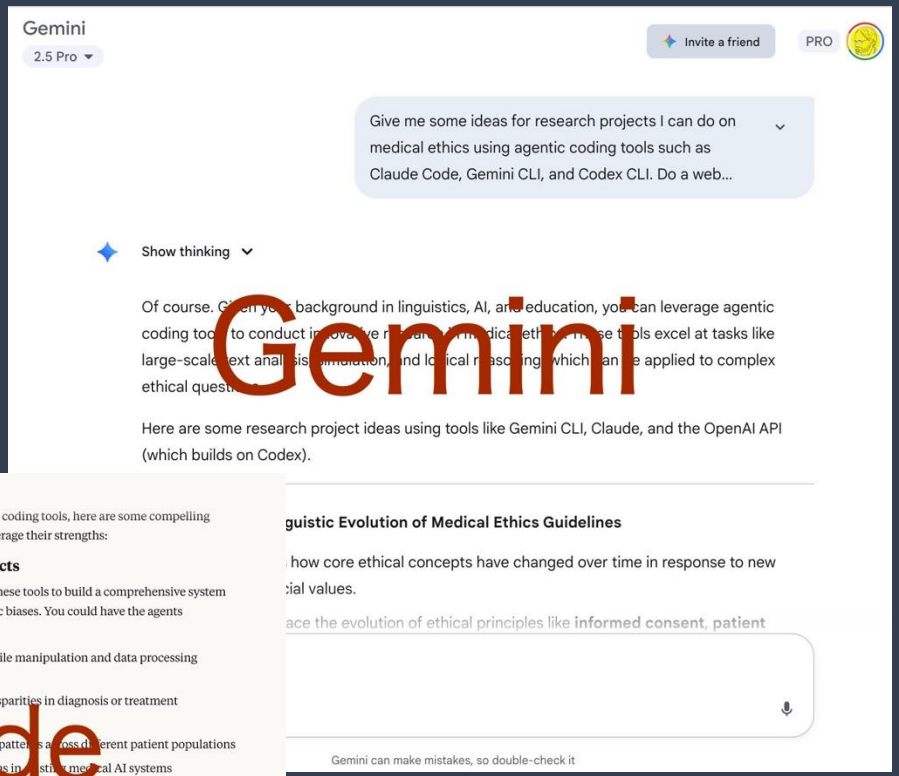
トム・ガリー

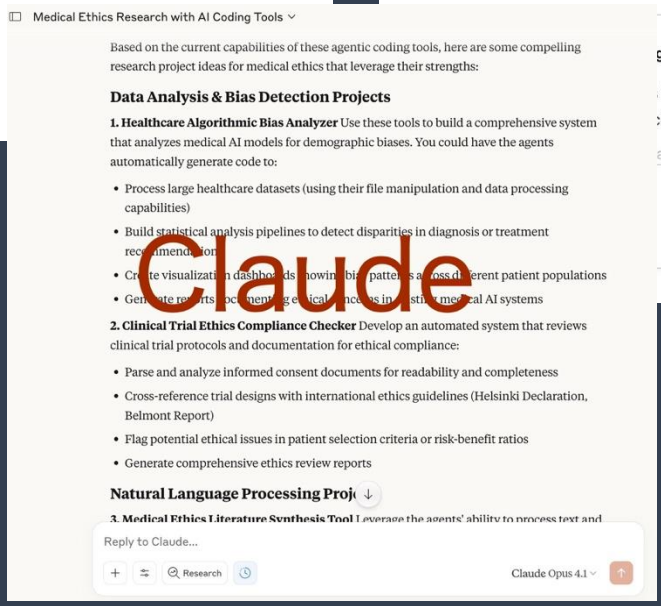
本日の内容

- 基本的なAIツール
- AIエージェント
- 研究評価への課題
- 何をすべきか？

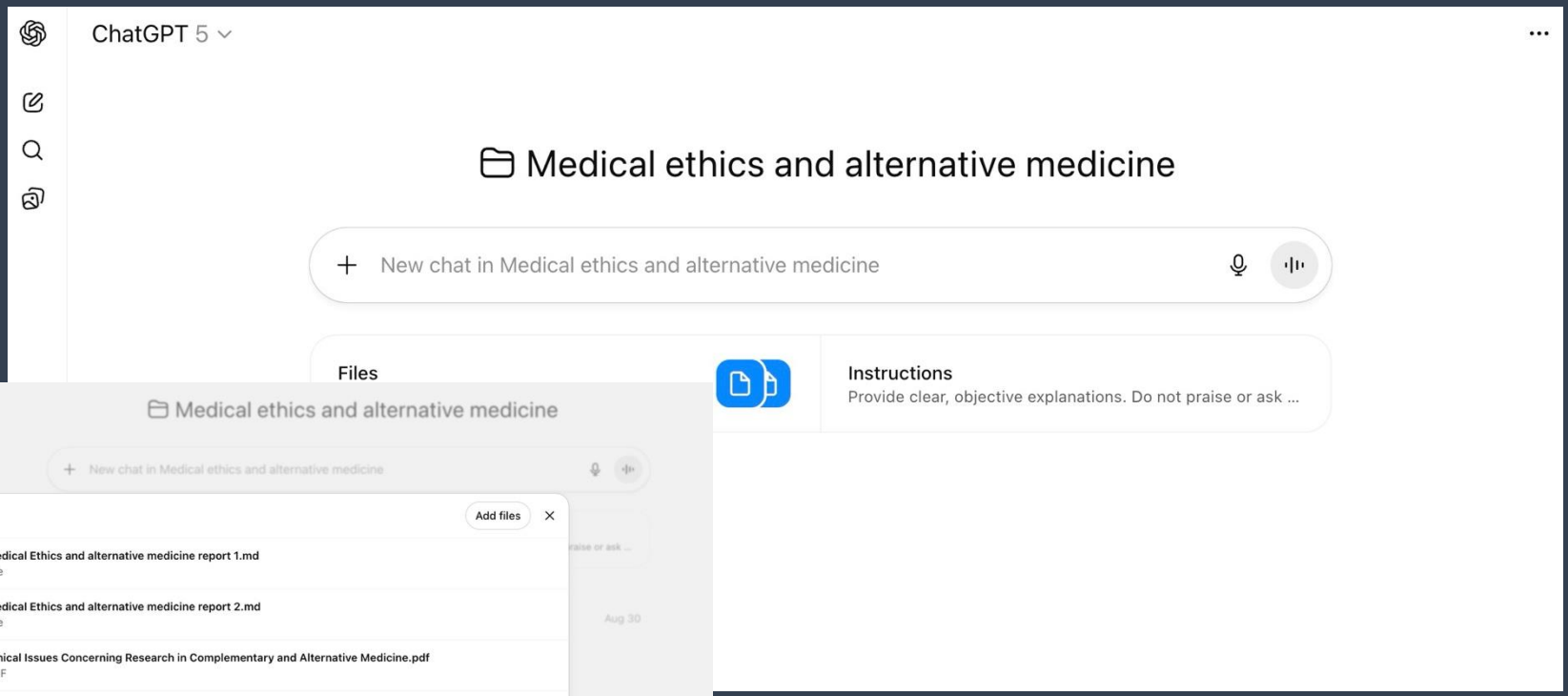
チャットのインターフェース



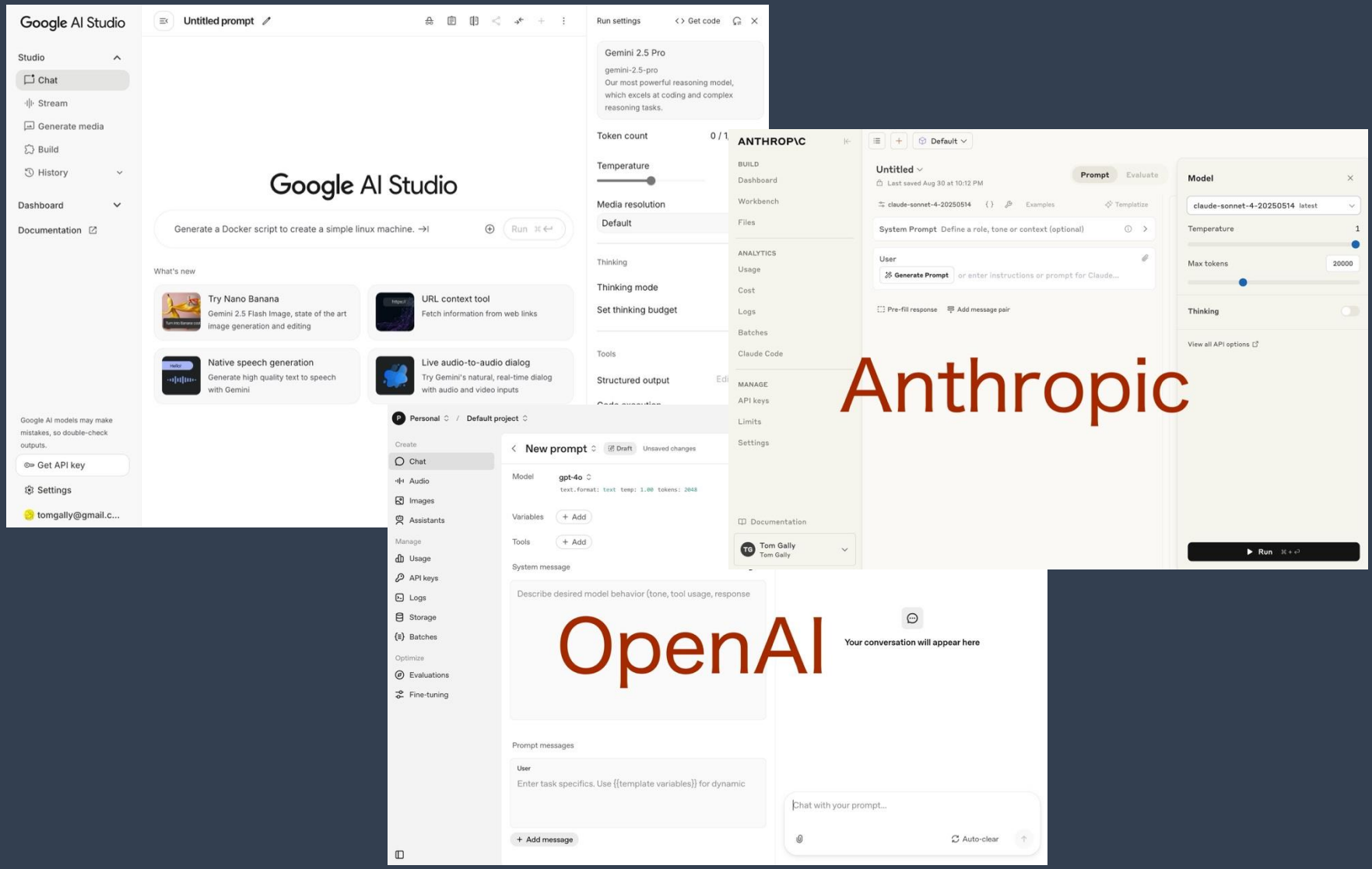




「プロジェクト」


ChatGPT 5
Medical ethics and alternative medicine
New chat in Medical ethics and alternative medicine
Files
Instructions
Provide clear, objective explanations. Do not praise or ask ...
Medical ethics and alternative medicine
New chat in Medical ethics and alternative medicine
Files
Add files
Medical Ethics and alternative medicine report 1.md
File
Medical Ethics and alternative medicine report 2.md
File
Ethical Issues Concerning Research in Complementary and Alternative Medicine.pdf
PDF
Physicians' Ethical Obligations Regarding Alternative Medicine.pdf
PDF
关于医学伦理学案例教学的内涵与外延.pdf
PDF
The J of Law Medi Ethics - 2007 - Sade - Complementary and Alternative Medicine Foundations Ethics and Law
PDF
보완대체의학의윤리적성찰.pdf
PDF
Complementary and alternative medicine.pdf
PDF
Medizinethik und Komplementärmedizin.pdf
PDF

ダッシュボードとAPI



基本的なAIツールでできること

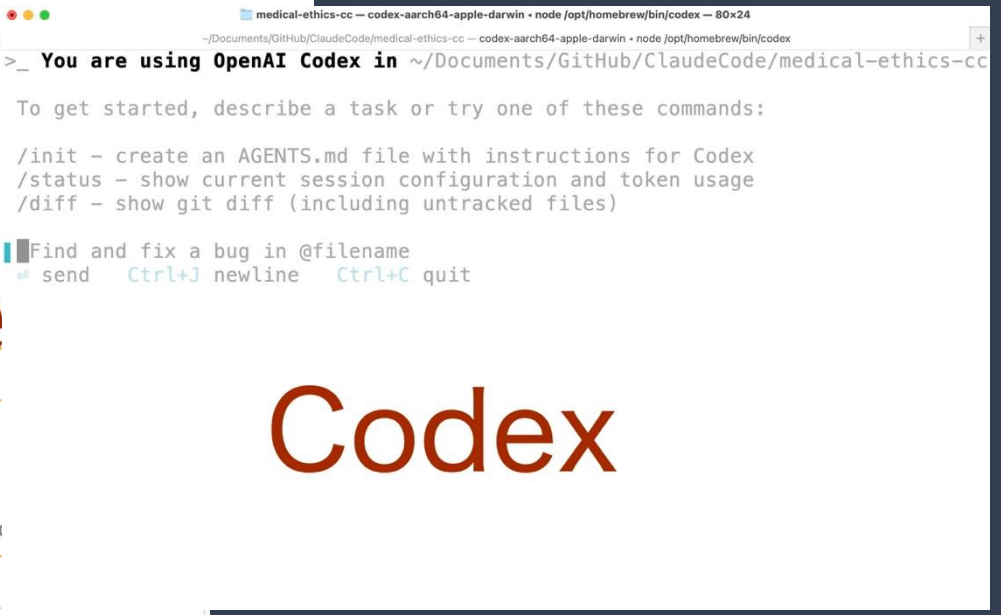
- 研究課題や方法のアイデアを得る
- 新しい分野への理解を深める
- 言語間の翻訳
- 文章を改善する
- 簡単なプログラムやマクロを書く
- LLMの長所と短所を理解する

AIエージェント

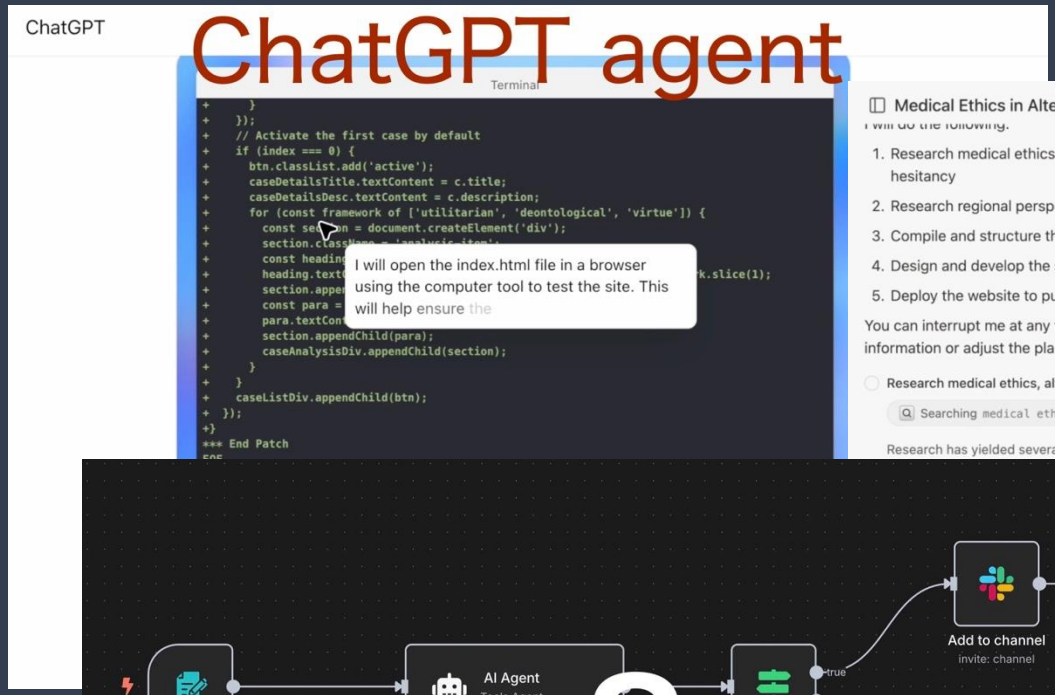
「AIエージェントとは、大規模言語
モデルが自身のプロセスやツール使用を
動的に管理し、タスクの達成方法を
調整し続けるシステムである」

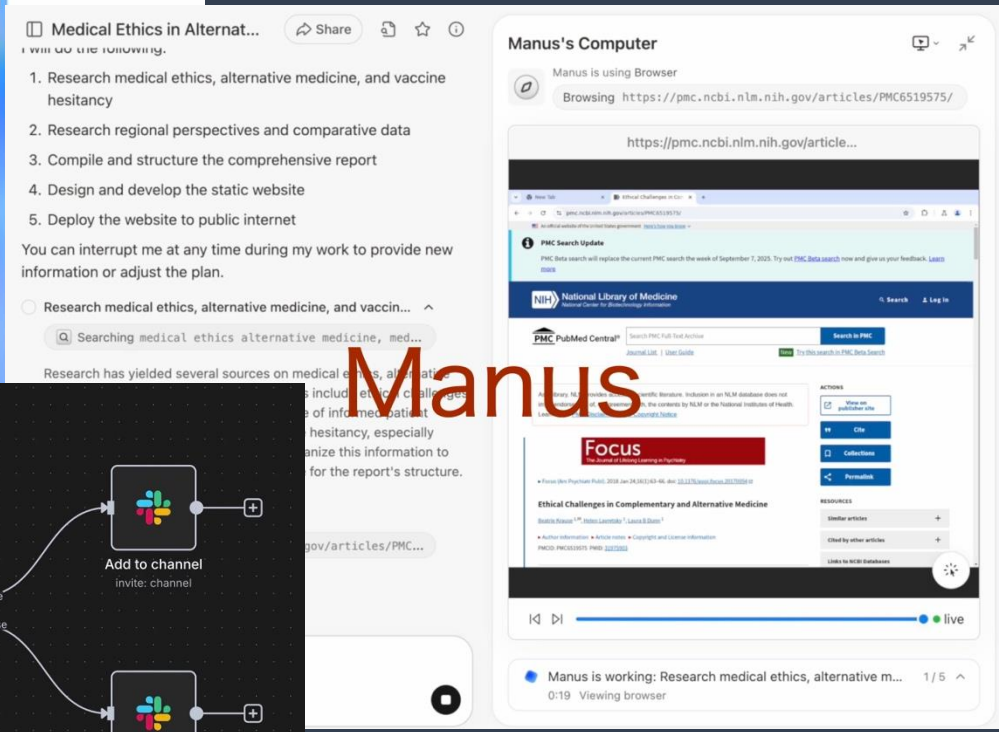
Anthropic, “Building Effective Agents”
2024年12月19日

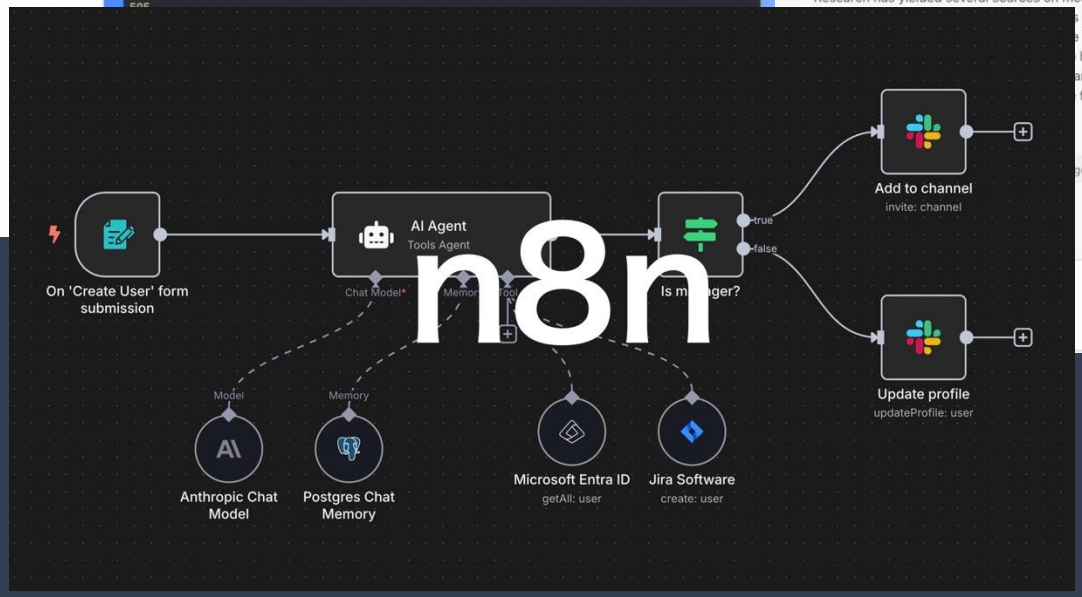
コマンドラインインターフェース（CLI）



ブラウザ統合型エージェント







エージェントができること

- ファイルの読み書き
- ウェブから情報を収集
- Pythonなどの言語でプログラムを書いて実行
- サーバーを起動・実行
- 必要なツールをダウンロード・インストール
- 複雑なタスクを計画して段階的に実行
- 複数のサブエージェント間でタスクを調整
- 大規模なコードベースを管理・修正
- ユーザーの許可があれば上記の作業を自律的に行う（サンドボックス環境が望ましい！）

エージェントで始める方法

1. ChatGPT、Claude、Geminiなどで「プロジェクト」を作成する
2. 研究分野に関連する論文などをアップロードする
3. Claude Code、ChatGPT agent、Manusなど、使用予定のエージェントを説明する文書をアップロードする。（これらのエージェントは新しいため、LLMにはおそらくその機能の詳細を提供する必要がある）
4. 次のようなプロンプトを試してみる：「私は【トピック】に取り組んでいる研究者だ。このプロジェクトには私が興味を持っている論文などが含まれている。また、【エージェント名】に関する情報も含まれている。私の研究にこのエージェントを使用できる20の方法を提案し、それらの目的で【エージェント名】を使用する詳細な手順を 教えて、その研究に使用できるプロンプト例を書いてください。」

現在のエージェントの問題点

- 自律的に作業できる一方で、時に脱線することがある。また、忘れっぽく、欺瞞的になることもある。人間の監視と指導が必要。うまくいかない場合は、別のプロンプトで再試行してください。
- エージェントはLLMに基づいているため、ハルシネーションを起こすことがある。複数のLLMが互いにチェックして修正するワークフローを設定することで、ハルシネーションを軽減できる場合がある。（複数のLLMを使用するにはOpenRouterが便利）
- エージェントはまだ初期段階にあり、急速に進歩している。使い方の専門家はいない。創造的で実験的に取り組んでください。
- エージェントには、プロンプトインジェクションなどのセキュリティ問題があり、効果的な対策はまだ開発されていない。

エージェントと研究評価

- 7月に東広島キャンパスでの講演で説明したように、一部の分野では、AIエージェントに研究を計画し、調査し、出版可能な論文を書かせることが可能だ。
- 世界中で多くの人々が、人間の知識を進歩させるため、または自分の業績を水増しするために、自律的または半自律的な研究エージェントを開発しようとしている。
- 人間がどの程度関与したか不明瞭な研究論文が大量に生み出される時代が目前に迫っている。

- 長年にわたり、研究を行うインセンティブは定量的評価によって歪められてきた（掲載数、引用数、インパクトファクターなど）。
- これらの歪んだインセンティブは、盗作、ずさんな方法論、研究不正など多くの問題を引き起こし、科学と科学者に対する社会的な信頼の低下をもたらしている。

研究の中核的な知的作業の多くが
AIによって行われるようになったとき、
研究と研究者はどのように
評価されるべきか？

「研究評価に関するサンフランシスコ宣言」（DORA）の抜粋

「個々の科学者の貢献を評価したり、採用、昇進、資金提供の決定を行う際に、ジャーナルインパクトファクターなどのジャーナルベースの指標を個々の研究論文の品質の代替指標として使用しないでください。」

「論文の科学的内容が出版基準や論文が掲載された雑誌のアイデンティティよりもはるかに重要であること」

「研究評価の目的で、研究出版物に加えて、すべての研究成果（データセットとソフトウェアを含む）の価値と影響を考慮し、政策と実践への影響など、研究の影響の質的指標を含む幅広い影響測定基準を考慮します」

「資金、雇用、終身在職権、昇進などを決定する委員会に参加する場合は、出版指標ではなく科学的内容に基づいて評価を行います。」

DORAへの日本の署名機関の例

- 日本学術振興協会
- 日本生物科学会連合（SEIKAREN）
- 国立研究開発法人医薬基盤・健康・栄養研究所
- 理化学研究所
- 東京大学※

※これまでにDORAに署名した唯一の日本の大学

DORA自体は、AI主導の研究が引き起こす課題に対処するには十分ではありません。しかし、私たちがこれらの非常に困難で重要な問題にどのように立ち向かうべきかのモデルになります。

AIの進歩は科学研究に大きな課題をもたらしています。私たちは皆、これらの進歩を把握した上、科学界がどのように対応すべきかについて、あらゆるレベルで積極的に議論する必要があります。