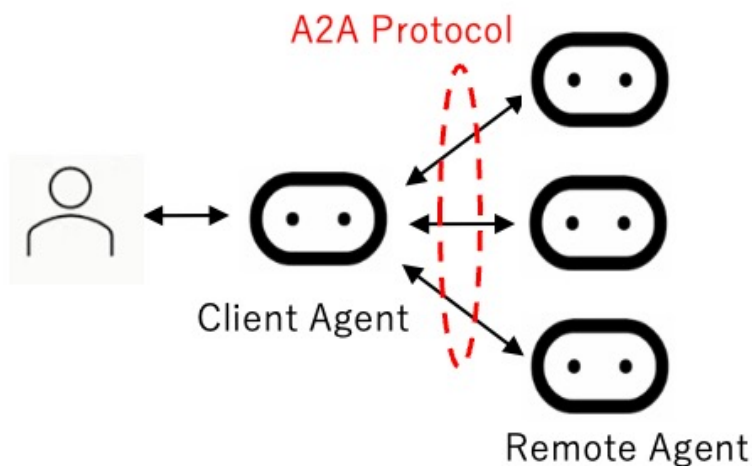


MCPとA2Aプロトコル

～ AIの開発史から知的協働作業の自動化まで～



2026年 3月

A2A研究所 田代 務

目次

| | | |
|----------|-------------------------------|----|
| 1 | AI開発の歴史 | |
| | 最初のAIブーム | 3 |
| | ディープラーニングの時代へ | 4 |
| | アテンションメカニズム | 5 |
| | AIの普及を牽引する事業者 | 6 |
| 2 | 生成AI | |
| | LLM（大規模言語モデル） | 7 |
| | 主な生成AIサービス | 8 |
| | オープンソースの生成AIサービス | 9 |
| | 画像生成AI | 10 |
| 3 | AIによる知的作業の効率化 | |
| | AIエージェント | 11 |
| | MCP (Model Context Protocol) | 12 |
| | MCPでのストリーミング | 13 |
| | A2A (Agent to Agent Protocol) | 14 |
| | A2Aプロトコルを使用する構成要素 | 15 |
| | A2Aプロトコルの賛同メンバー | 16 |
| 4 | A2Aプロトコルによる協働作業 | |
| | ヘテロジニアスなマルチエージェント | 17 |
| | エージェント・カード | 18 |
| | 非同期通信モード | 19 |
| | A2Aプロトコル使用の協働作業例-1 | 20 |
| | A2Aプロトコル使用の協働作業例-2 | 21 |

最初のAIブームは1960年代に始まった

AI(人工知能)という言葉は1956年に米国ダートマス大学で開催のワークショップにて同大学 ジョン・マッカーシー教授が「学習や知能の全機能は正確に機械でシミュレートできるので、これをArtificial Intelligenceと呼びたい」と提案したことに由来するようだ。その後、AIの長い歴史が始まることになった。

第1次AIブーム：「推論と探索の時代」

1950年代から70年代までは、AI応用は最高性能の計算機でさえ幾何定理の証明や英会話の学習程度の極めて特定領域での問題に限られた。しかし、そうした計算機の知的能力は驚異的と捉えられ、AI研究者の間にも将来の計算機能力進歩に楽観的な予測が広がった。なお、日本では1967年に甘利俊一が現在のAIの基礎となる「多層パーセプトロンの確率的勾配降下法」に初めて成功した。その後、ヒントン等が誤差逆伝搬法として再発見している。

第2次AIブーム：「知識工学の時代」

1980年代、AIの一種であるエキスパートシステムが多くの企業で導入された。特定領域での質問の応答や問題を解くシステムであり、分光計測定結果からの化合物の特定や血液疾患の診断など幾つかの成功例があったものの広範な応用には至らなかった。1981年、通産省は第5世代コンピュータプロジェクトを開始し、自然言語での人間との対話、機械翻訳、画像認識などで人間のような推論を目指したが成功しなかった。

多層パーセプトロン

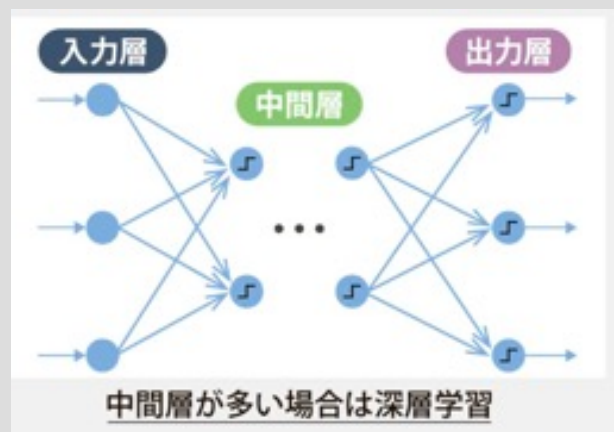
◆誤差逆伝搬法

出力値と正解の差(誤差)を出力層からさかのぼって誤差が小さくなるようにネットワークの重み W を調整する学習手法

◆確率的勾配降下法

誤差逆伝搬法の具体的なやり方の一種。誤差が小さくなる方向(勾配)を見つけその方向にネットワークの重み w を調整する。学習データの一部をランダムに選ぶことにより、勾配方向にランダム性を取り入れ最適な解を探索できるようになる。

【パーセプトロン中間層のイメージ】



電気新聞記事(2023.3.27) 堀 修 より

現在のAI心臓部の原型とも言える多層パーセプトロンは半世紀も前に考案されていた。



1. AI開発の歴史

ディープラーニングの時代へ

2024年のノーベル物理学賞と化学賞はいずれもAI関係に

第3次AIブーム：「ディープラーニングの時代」

2006年にジェフリー・ヒントンが人手を介さず特徴量を抽出できるディープラーニングを発明し、大きなブレークスルーとなった。2012年にはディープラーニングを用いたYouTube画像からの猫の認識にGoogleが成功したと発表し、世間の注目が集まり始める。

第4次AIブーム：「LLMと生成AI」

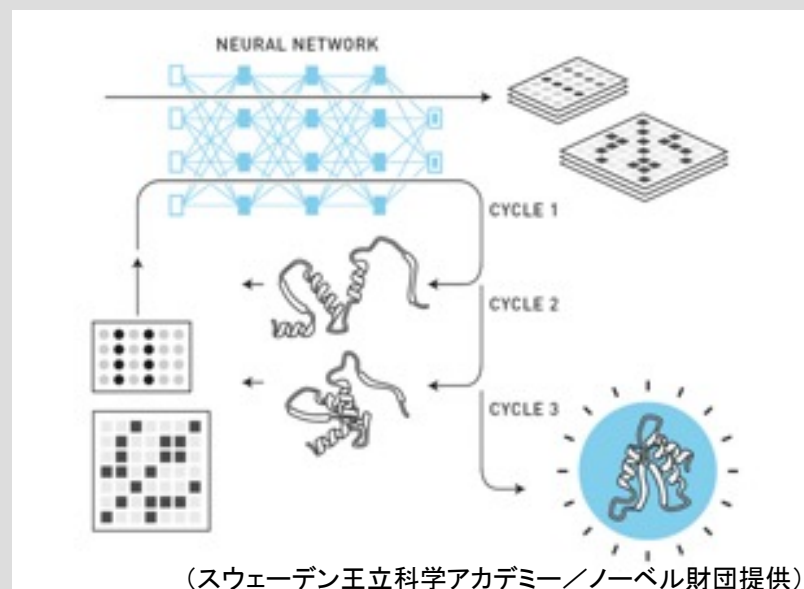
2020年、OpenAIが自動言語処理プログラムGPT-3を開発する。ニューラルネットワークと自己注意機構を基盤としたトランスフォーマーと大規模言語モデル(LLM)を用いた。このLLMの学習・推論には膨大な浮動小数点の高速並列計算のため、当初はスーパーコンピュータが試みられた。その後、ゲーム端末向けのNVIDIA製画像処理装置(GPU)で対応できることが知られてから研究者での利用が急速に広がった。

2024年には、ジェフリー・ヒントン等が「ニューラルネットワークによる機械学習を可能にする基礎的な発見と発明」によるノーベル物理学賞を受賞する。

同年には、グーグルCEOや研究者等が「AIを用いたタンパク質の構造予測」によりノーベル化学賞を受賞し、生成AIが現代社会を支える重要基盤になったことを世界に強く印象づけた。

一方で、生成AIの悪用による偽情報の拡散などの影の部分も一段と強くなっている。

アミノ酸配列からタンパク質構造を予測する方式のイメージ図



2024年のノーベル物理学賞と化学賞はいずれもAIの実用化に関わる受賞だった。



1. AI開発の歴史

アテンションメカニズム

注意こそが必要とされる全てだ

AI技術の初期の応用で最も成果を上げた分野は機械翻訳だった。

2017年にGoogleの研究者達が発表した論文は、高性能の機械翻訳を達成した深層学習モデルTransformerに関するものであり、その核心部分を次のように表現した。
「Attention Is All You Need」

このTransformerは概略、次のような構成としている。

- 原文をトークン(ほぼ単語に相当)に分別し、トークン列(ベクトル)に変換する。
- 原文側、翻訳文側の両者に再帰型あるいは畳み込み型のNN(Neural network)を用いる。
- 両者のNNにアテンションメカニズムを適用する。

アテンションメカニズムとは文章中から抽出されたトークン列において、近傍トークン間の関連度(あるいは類似度)を0~1に数値化する過程を通じて重要なトークンに注目(注意)する機能である。

このモデルは並列化が可能であるので並列処理に適したGPUを用いるとトレーニング時間が大幅に短縮され、短時間で高い機械翻訳品質(BLEU: Bilingual Evaluation Understudy スコア)を達成できた。

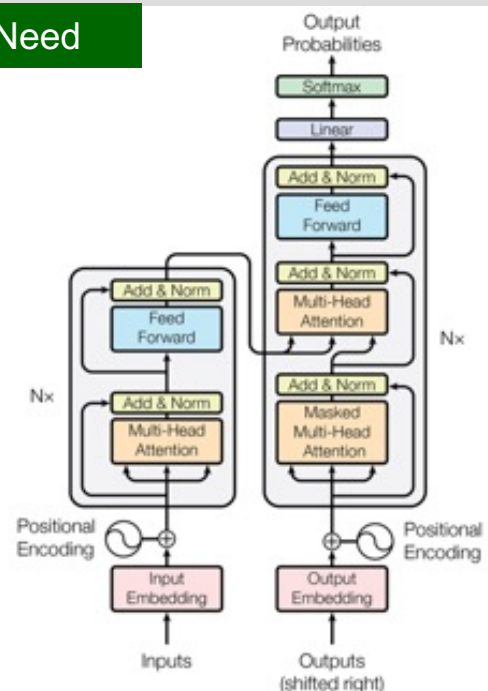
Attention Is All You Need

Google研究者による“Attention Is All You Need”と題する記念碑的論文の図1 The Transformer – model architecture より

図の左側は原文側のデコーダ、
右側は翻訳文側のエンコーダ

ニューラルネットワーク(Nx)内で、文章内のトークン系列間の関連性を行列で表現し、行列式の計算から関係性が高いとAttention という数値で求め行列化する。

例えば、「文章を要約してください」という文章の場合、「文章」「要約」というトークンが重要であると理解することになる。



画像や文章の特定部分に注目して見るAttention は人間が普段行っていることでもある。



1. AI開発の歴史

AIの普及を牽引する事業者

OpenAIとアンソロピック

近年、AIの動向に関する報道がマスコミを賑わしている。
 例えば、AIのソフトウェアコーディング能力は熟練技術者を凌ぐようになったので、もはや技術者はコーディングは行わず専らAIをチェックするだけになった等々。
 数年前までは、AIの能力が人間を超える所謂シンギュラリティ(技術的特異点)の到来は2040年代中頃と言われていたが、最近の予想では2030年代初め頃と大幅に早まるようである。

こうしたAIの破壊的革新を牽引する代表格はオープンAIとアンソロピックだろう。
 オープンAIは、人間からの数多な質問に瞬時にほぼ正確に答えるチャットGPTを瞬く間に全世界に普及させた。
 その後、テキストだけでなく画像や動画、音声や音楽、ソフトウェアプログラム等の生成に適した生成AIサービスを提供する事業者が続々登場し、教師やコンサルタントからクリエイター等の広範な知的労働者に深刻な影響を与えつつある。

一方、企業向けの各種業務の自動化を専門とするアンソロピックはSaaS(Software as a Service)事業者の強敵となると目されている。企業間ビジネス用のAIエージェント間プロトコル登場も相まって企業内や企業間のビジネスに大きな革新が生じるかもしれない。2026年2月には高性能AIが従来のサービス需要を奪うとの警戒感から関連ビジネスを展開する企業の株価が大幅下落する「アンソロピック・ショック」も起こっている。

オープンAIとアンソロピックの年表



オープンAI サム・アルトマン

| | |
|-------|--|
| 1985 | セントルイスで出生。母は皮膚科医でユダヤ人家系。 |
| 2005迄 | スタンフォード大で計算機科学を学ぶ |
| 2015 | イーロンマスクやピーターディール等とともに研究企業のOpenAIを設立し、CEOに就任 |
| 2023 | OpenAIを退任し、マイクロソフトに入社を発表。その後、OpenAIに復帰 |
| 2024 | OpenAIは非営利組織のOpenAI Foundationと営利組織のOpen Groupで活動中 |



アンソロピック ダリオ・アモディ

| | |
|------|---|
| | カリフォルニア大、スタンフォード大で物理学学士、プリンストン大で神経回路の研究で物理学博士 |
| 2014 | Baidu入社後、Googleへ転職 |
| 2016 | OpenAIに入社 |
| 2021 | 妹のダニエラやOpenAIの他の元メンバーとAnthropicを設立 |
| 2023 | 米国上院司法委員会で、兵器開発と管理にAIのもたらす危険性を警告 |
| 2023 | アモディ兄妹はTIMIE誌の「AI分野で最も影響力のある100人」に |

オープンAIと袂を分かったアンソロピックは企業内や企業間のビジネス革新に注力する



2. 生成AI

LLM（大規模言語モデル）

数百もの隠れ層を持つニューラルネットワークが一般的

LLM(Large Language Model: 大規模言語モデル)は膨大な量のテキストデータを深層学習(ディープラーニング)技術を用いて学習することで、言語を高度に理解・生成するAIモデルである。最も一般的なモデルは、入力層と出力層の間に数百もの隠れ層(Hidden Layer)がある多層NN(Nural Network)モデルである。

多層NNでの数億から数千億個のパラメータの各数値は、大量のデータ(Web上の記事、書籍、ニュース、会話記録等)による学習を通じて決定され、LLMは次の機能を備えることになる。

- ・質問への回答
- ・文章の要約
- ・機械翻訳
- ・プロンプト(入力)の続きの予測
- ・キーワードの抽出
- ・プログラムのバグチェック など

LLMはテキストの理解や生成に特化したモデルであるが、画像、映像、音声などのテキスト以外のコンテンツを自動生成する方式にも応用され、生成AIと総称されることが多い。

主なLLM

出典:Smiley社サイト等を元に作成

| モデル名 | 開発者 | 発表年 | パラメータ数 | 特徴 |
|-----------|-----------|------|---------|---|
| BERT | Google | 2018 | 3.4億 | データセットの規模拡大で精度を向上させた初期モデル |
| GPT-4 | OpenAI | 2023 | 非公開 | GPT-3に画像や音声などテキスト以外を学習 |
| PaLM | Google | 2022 | 5,400億 | Transformerのパラメータ数を大幅拡大 |
| LLaMA | Meta | 2023 | 70~650億 | GPT-3と同等性能を少ないパラメータ数で実現 GitHub上にオープンソースで公開 |
| Claude | Anthropic | 2023 | 非公開 | GPT-2,3開発技術者によるモデル |
| Gemini1.5 | Google | 2024 | 非公開 | Google DeepMind開発のマルチモーダルモデル |
| OpenELM | Apple | 2024 | 約10億 | 複数層を作り、各層内パラメータを効率的に配置 |
| GPT-5 | OpenAI | 2025 | 非公開 | コーディング・数学、文章作成等の分野で利用可能 |
| Tsuzumi-2 | NTT | 2025 | 6~70億 | 軽量(少ないパラメータ数)で日本語と英語に対応 |
| cotomi | NEC | 2024 | 130億 | 軽量で日本語に対応 |
| ELYZA | 東大 | 2025 | 700億 | 松尾研究室発AIカンパニー*による日本語特化型 |

*) KDDI子会社

現在主流のLLMは数百から数千億個ものパラメータを有している。



2. 生成AI

主な生成AIサービス

文章生成から動画や音楽の生成も

学習データからテキスト、画像、音楽、動画等を自動作成するAI技術である生成(Generative) AIが急速に進化している。

例えばオープンAIが提供するチャットGPTの最新版(GPT-5)では、テキストに加え画像や音声の生成や高度な推論が可能になっている。適切なコンテキスト(背景)や具体的な要求をプロンプト入力することでデザイン性に優れたWebサイトやゲームも作成できる。

生成AIサービスの提供企業にはAI専門ベンチャーのOpenAIやAnthropic, ビッグテックのGoogle(親会社はAlphabet)やMicrosoft, デザインやソフトウェアの巨人Adobe等のほか、画像生成に特化したAIをオープンソースで提供するStable Diffusion, 同じくオープンソースで楽曲生成するVOICEVOX等々、業界はまさに百花繚乱状態にある。

文章、画像、動画、音声・音楽の4種類の生成物では次のような利用が想定されている。

- ・文章生成 報告書やメールの作成、文章の要約、翻訳、議事録作成
- ・画像生成 SNS投稿や広告用の画像の作成
- ・動画生成 プロモーションビデオ、教育向けコンテンツの制作
- ・音声・音楽生成 ナレーションや音声ガイドの作成、音楽制作

| 主な生成AIサービス | | | | |
|-------------------|------------------|-----------|----------------------------------|---------------------------------------|
| 富士フィルム社サイト資料を元に作成 | | | | |
| 生成物 | サービス名称 | 開発/提供者 | 月額料金(注) | 特徴 |
| 文章 | ChatGPT | OpenAI | Plus:20 \$, Pro: 200 \$ | 汎用性高い、画像・音声も可 GPT-5ではWeb, アプリ、ゲームも |
| | Gemini | Google | Pro: 2900円 | 最新情報の検索や長文処理に優れる |
| | Claude | Anthropic | プロ:17 \$, マックス: 100 \$ | 倫理的に正確、日本語が自然 |
| | Copilot | Microsoft | Personal: 2130円 Premium:3200円 | Officeツールと連携可 |
| 画像 | Stable Diffusion | オープンソース | プロ:7 \$, マックス: 14 \$ | オンプレミスでモデルを自由に調整可 |
| | Adobe Firefly | Adobe | Standard: 1580円 Pro:4780円 | Adobeツールと連携可、安全性を重視 |
| 動画 | Sora | OpenAI | ChatGPT Plusに加入 | 高いプロンプトの理解力で映像表現 |
| | Runway | | 標準:15 \$, プロ:35 \$ | 画像からの動画生成 |
| 音声・音楽 | VOICEVOX | オープンソース | 無料 | 日本発の音声合成ソフト |
| | Suno AI | | プロ:1500円 | プロンプトから楽曲を数分で生成 |

(注) 上は個人むけの代表的なプラン、殆どのサービスに無料版あり

あらゆる形式の知的生産物を効率的に生成できる生成AIサービスが次々に現れている

文章生成から動画や音楽の生成も

オープンソースAIはソースコードが公開され誰でも自由に使用、変更、配布ができるAIである。

GitHub等のソフトウェア開発者向けプラットフォームではAIアルゴリズム、事前トレーニングモデル、データセットが公開されている。そこで次のようなライブラリを用いて、ボランティア愛好家のコミュニティは既存の諸作品をベースに実用的なAIソリューションを開発している。

- ・TensorFlow: 数式やニューラルネットワーク(NN)の処理を高速計算
- ・PyTorch: Pythonベースの直感的な操作性がある機械学習ライブラリ
- ・Scikit-learn: 機械学習と予測データ分析用のPythonライブラリ
- ・OpenCV: リアルタイムの画像・動画処理用の各種アルゴリズムを提供

企業独自のAIモデル(クローズドソース)に比べてカスタマイズの自由度が高く、ローカル環境で実行できるので低コストでプライバシー確保にも優れるが、一方でライセンス形態によっては責任の所在や持続可能性の問題が生じたり、高性能のGPUが必要な場合がある。

一般ユーザの使用のみであれば、画像や音楽を生成する次のようなサービスもある。

- ・画像生成AIのStable Diffusion
- ・歌声合成AIのVOICEBOX

オープンソースの生成AIサービス例



上: Stable Diffusionにて「サングラスをかけた猫」のプロンプトへの返答
GRUSOROBANサイト <https://soroban.highreso.jp/article/article-034>
右: VOICEVOX 青山龍星による無料のテキスト読み上げ
VOICEVOXサイト <https://voicevox.hiroshiba.jp> より



初心者でもオープンソース生成AIサービスで比較的容易に画像や音楽が生成できる。

拡散モデルが代表的な方式

画像生成AIは、極めて多数(数億～数十億枚)の画像とテキストのペアをディープラーニングで事前学習させておき、ユーザから入力されたテキストから学習済みデータを基に特徴を探し出し一致する特徴から画像を生成する仕組みである。

例えば、「ネコ、昼寝」と入力すると昼寝をしている猫の写真やイラストが生成されるようなもの。様々な提供サービスから、実写・アニメ・3Dモデル等の生成画像の種類を選択できる。

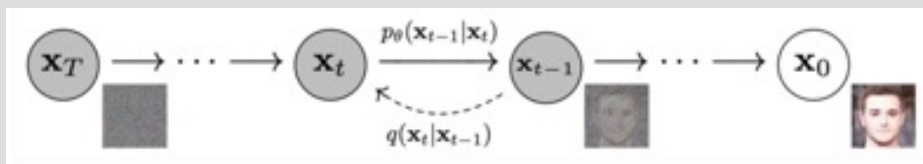
この画像生成AIには様々な手法があるようだ。

その代表格は拡散モデル(Diffusion Model)であり、オープンソースで画像データ生成サービスを提供するStable Diffusionが使用している。

拡散モデルの第1ステップでは要求画像に類似した実画像データに少しずつノイズを足してノイズ画像を作成する。次に、このノイズ画像から逆にノイズを徐々に除去する過程で、AIが過去に学習した大量の画像やプロンプトの指示文を元に形や色を補完し、リアルで自然な要求画像に仕上げていく。

画像生成AIを用いて誰でも簡単に高精度の画像が作成できるが関係法律が未整備のため、利用目的を誤ると著作権や知的財産権を侵害する恐れがあるので注意が必要である。例えば、既存のキャラクターやイラストに酷似した画像を生成し利用する場合など。

拡散モデルでの画像生成



"Denoising Diffusion Probabilistic Models" <https://arxiv.org/pdf/2006.11239.pdf> より



例えば、左の類似画像を元に、右の要求画像を生成する

類似する実画像にノイズを加えた後、そのノイズを逆に除去していく過程でプロンプトの指示等の情報をもとに形や色を徐々に補正していくことで要求画像を生成する。

人に代わって業務を自動実行するソフトウェア

AIエージェントは人が細かな指示を与えなくても業務の目標を自ら計画し、必要情報を探し、様々なツールを使い分けてタスクを完遂する自律型ソフトウェアである。

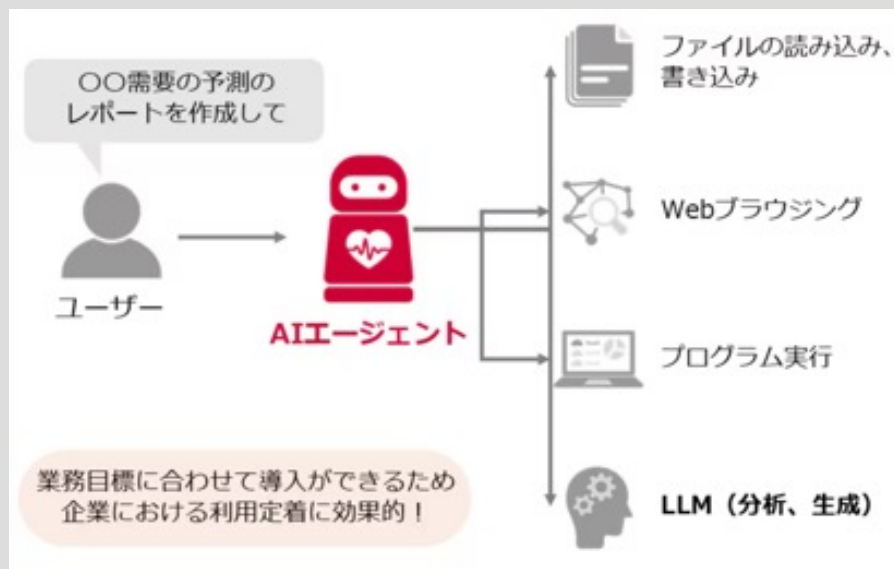
例えば、コールセンターのAIエージェントは人に代わって次のような作業を自動実行できるだろう。

- ・ 顧客との会話内容に基づき、FAQをリアルタイムで検索する
- ・ 顧客との通話後に会話を要約し、必要に応じて社内共有用の知識ベースを作成する

企業内ではこうした問い合わせ対応のほか、経理、保守などの現場業務や複数部門間のつなぎ役の仕事もAIエージェントに置き換えやすい。

社内業務でのAIエージェントは下図のようにWeb検索、ファイルの入出力、Web検索、プログラムの実行など複数のツールを適宜利用してユーザに代わって自動的にタスクを実行完了する。

AIエージェント利用イメージ



出典:ドコモビジネスのサイトより

コールセンターでは既に人に代わって問い合わせに対応し、会話の要約も作成することも



3. AIによる知的作業の効率化

MCP (Model Context Protocol)

アンソロピックが発表

MCP(Model Context Protocol)は、大規模言語モデル(LLM)と外部システムを接続するための標準規格(プロトコル)であり米国Anthropic社が2024年11月に発表した。

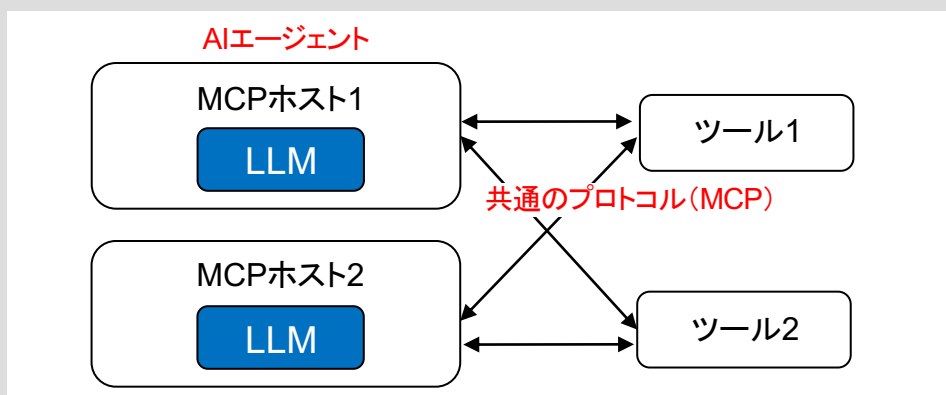
LLMを内蔵するAIエージェント等の業務向けAIアプリでは外部ツールから必要なコンテキスト*)を適切に供給する仕組みが重要である。このコンテキスト供給を標準化するプロトコルがMCP (Model Context Protocol)である。

*)コンテキスト: LLMの正しい推論と応答に必要な補足情報のこと。直前までの会話履歴、指示の背景や目的、時間・場所・組織等の状況情報、ユーザの過去の行動や嗜好等がある。

それまで、OpenAI, Google等が開発した方式は独自仕様のため異なる実装が必要だったが、MCPによって複数のLLMが利用可能となるシステムの構築が可能になった。そこで、異なる“デバイス”(LLM)にも同じ“ケーブル”(MCP)で接続できるので「AI用のUSB—Cポート」に例えられることもある。

また、Anthropicはプロトコルに加え、Python等の言語に対応したソフトウェア開発キット(SDK)を同時発表した。これによって多様なAIアプリとデータソース間が連携できるようになり、AIを単なる会話アプリから知的作業パートナーへ進化させつつある。

MCP (Model Context Protocol)



MCPホストはLLMを搭載したChatGPTやClaudeなどのアプリケーション本体
一方、ツールには、次がある

- Web検索: Google検索など
- 社内DBの照会: GoogleドライブやGmailなど
- Web API呼び出し: 画面上のクリック等で外部サービスや社内システムと連携してデータ取得や処理を行う

MCPによって複数のLLMが利用可能となるシステムの構築が可能になった



3. AIによる知的作業の効率化

MCPでのストリーミング

大量データの送信にはデータを小分けにして複数回伝送する

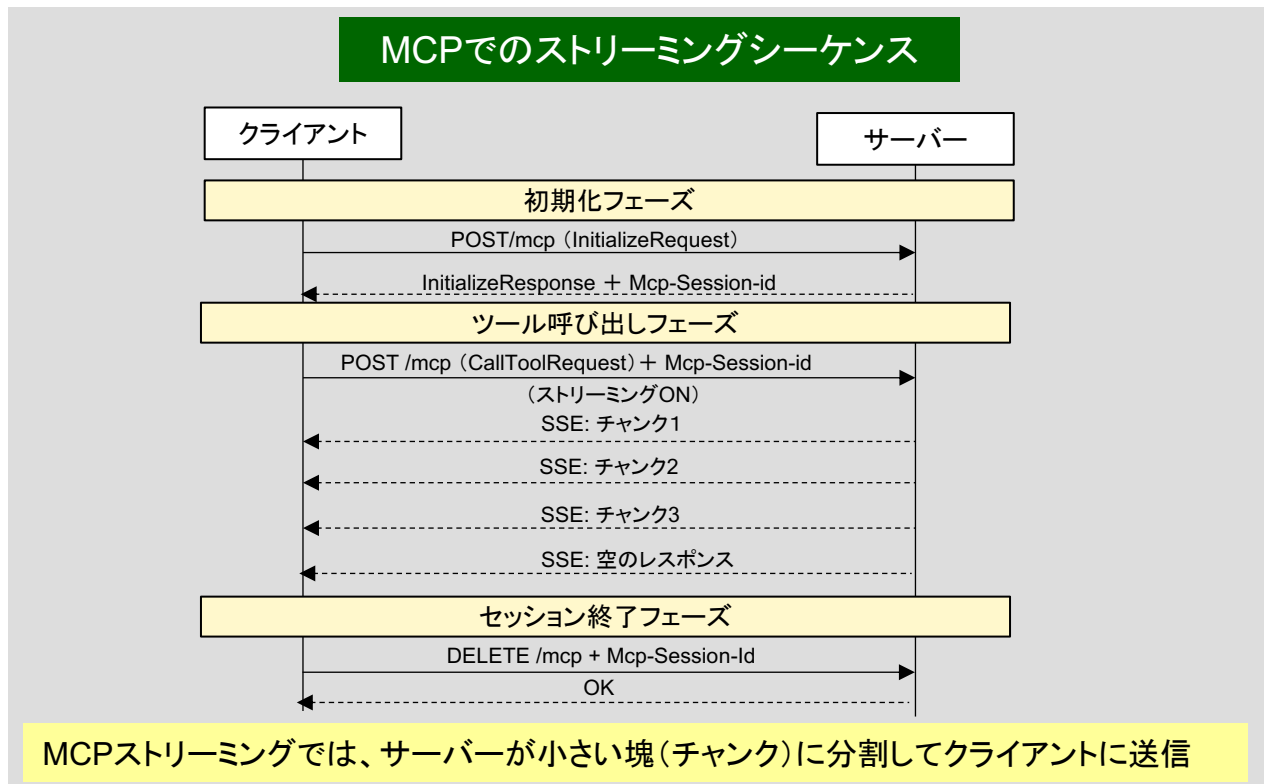
MCP(Model Text Protocol)では、Webブラウザとサーバ間での要求/応答型の通信手順(プロトコル)であるHTTP(Hyper Text Transfer Protocol)に加え、動画や楽曲でお馴染みのストリーミング機能が搭載された。

多くの場合、HTTPのリクエスト・レスポンスモデルで十分だったが、リアルタイム性が必要な場合や、大量のデータ処理のためクライアントに長時間の待合せが生じる場合は不都合だった。

そこでMCPではストリーミング機能を標準搭載し、更に、処理の進捗状況の通知や中間結果の配信などの双方向通信を加えることで対話的なアプリケーション構築を可能にしている。

下図はMCPでのストリーミングシーケンスの例であり、サーバーはデータを小さな塊(チャンク)に分割し、その各々をSSE(Server-Sent Event)としてクライアントに順次送信している。最終のSSEは「送信完了」を伝える合図である。

多くの生成AIのサーバーではクライアントとの接続状態を維持したまま、少しずつデータをプッシュ送信することでリアルタイムでの回答通知を行うようにしている。





3. AIによる知的作業の効率化

A2A (Agent to Agent Protocol)

2025年にGoogleが開発

複数の企業等のAIエージェント同士が連携して複雑なタスクを実行する世界が訪れている。A2A(Agent to Agent Protocol)は、AIエージェント間のプロトコルとして2025年にGoogleが開発し、その後、Linux Foundationに寄贈された。

A2Aプロトコルには次の特徴がある。

- ・HTTPとJSON-RPC ベースの簡単な設計
- ・ストリーミングによるリアルタイム通信に対応
- ・プッシュ通知による非同期タスクの管理
- ・企業向けの厳格なセキュリティ機能

また、A2Aプロトコルには次の構成要素がある。

- ・クライアントエージェント： ユーザ(人)の代理として他のエージェントに作業を依頼
- ・リモートエージェント： HTTPエンドポイントの公開などで依頼された作業を処理
- ・エージェントカード： エージェント能力、認証方法、エンドポイント等が記載された身元証明書

A2Aプロトコルでは相手エージェントにリクエストを送ると、レスポンスとしてMessageまたはTaskが返送される。Messageは単発のやり取りで完結する応答である。一方、Taskには状態があり submitted → working → completed の状態変化で処理の進行状態を管理できる。

A2AプロトコルのMessageとTaskの比較

| 判断項目 | Message | Task |
|--------|----------|---------------|
| 処理時間 | 短時間(秒単位) | 長時間(分～時間) |
| 状態管理 | 不要 | 要 |
| 進捗確認 | 不要 | 可能 |
| キャンセル | 不可 | 可能 |
| 成果物管理 | レスポンスのみ | Artifactとして保存 |
| リソース使用 | 軽い | 重い |

Messageの例

- ・「こんにちは」への挨拶
- ・簡単な質問への回答
- ・計算結果の通知
- ・状態の確認

Taskの例

- ・レポートや文書の作成
- ・大量データの分析
- ・画像や動画の作成
- ・複数ステップの処理

Messageは一问一答のような応答、Taskは複雑な処理や状態管理が必要な作業に



3. AIによる知的作業の効率化

A2Aプロトコルを使用する構成要素

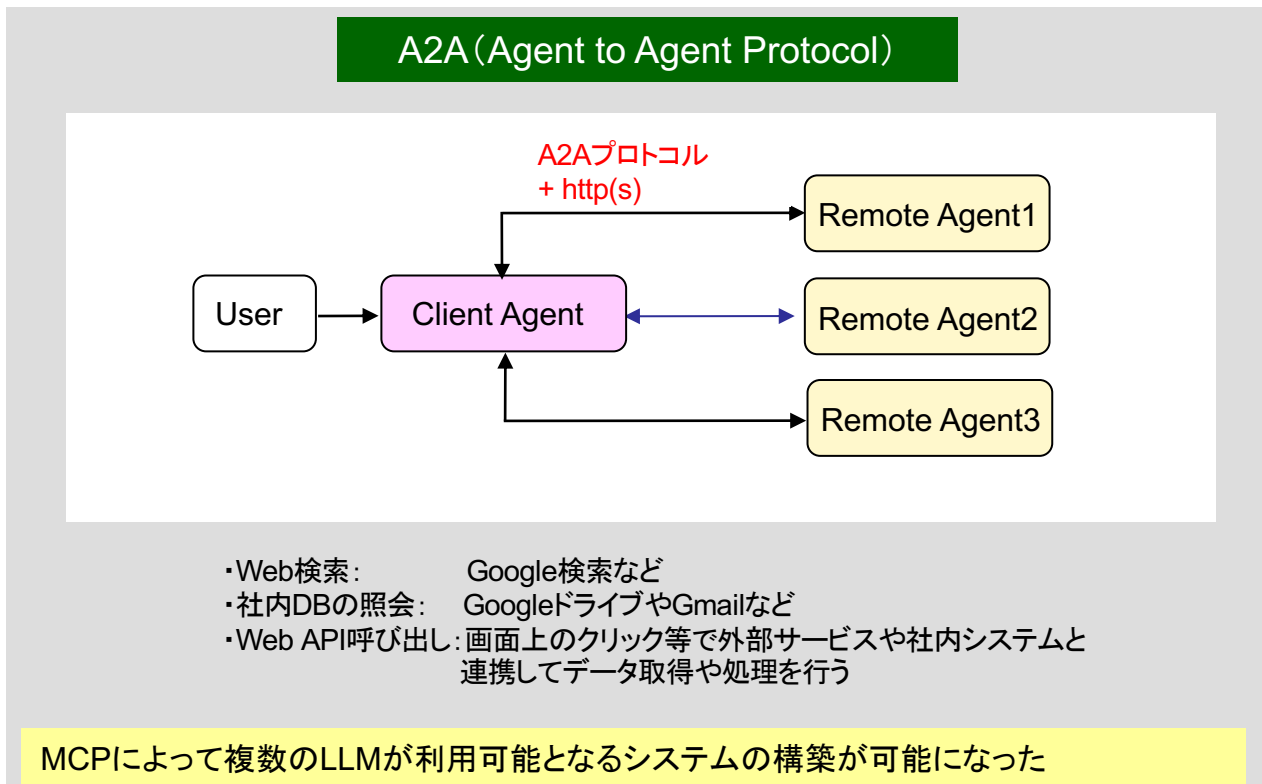
クライアント・サーバーモデルを採用

A2A(Agent to Agent Protocol)は、クライアント側のタスクの実行を1つ又は複数の外部エージェントに依頼して協働作業を自動的に行う場合に使用するプロトコルである。

MCPは主に企業内にて、LLMを内蔵するAIアプリケーションがデータソースやツール等の外部サービスやリソースを利用する際に用いられるのに対して、A2Aは外部のAIエージェントとの間の共同作業に使用される。

利用シーンは例えば、ユーザが旅行計画を作成するといった場合である。顧客からこれを依頼された旅行代理店はホテル等の宿泊施設、鉄道や航空等の交通機関の予約等はこれらに精通した各々のエージェントに作業を依頼して全体の旅程や経費を作成するといったものである。

データのやりとりには通常、Webブラウザとサーバー間でファイルや画像などのコンテンツを転送するための通信プロトコルであるHTTP(Hypertext Transfer Protocol)や、その通信を暗号化したHTTPS(HTTP Secure)が使用される。





3. AIによる知的作業の効率化

A2Aプロトコルの賛同メンバー

2025年にGoogleが開発

AIエージェント間のプロトコルA2Aは Anthropic が開発したMCPを補完するプロトコルとして、2025年にGoogleが専門知識を活かして開発されたので当初から多くの賛同者を集めている。

A2Aは大規模なマルチエージェント システム導入の課題に対応するよう設計されており、とりわけ多様なプラットフォームやクラウド環境でのエージェント管理を標準化できる相互運用性が注目されているようだ。

GoogleはA2Aプロトコルの公開にあたって次のように述べている。
”A2Aは開発者に当該プロトコル準拠の他のエージェントと接続可能なエージェントを構築する機能を提供し、多様なプロバイダーのエージェントを組合せる柔軟性をユーザーに提供します。”

Google Cloudサイトにて2025年4月に公表されたA2Aプロトコルの初期賛同メンバーは米国本拠の約60社であるが、次のサイトによると2026年3月時点では他国からの参加数が増加し合計で約120社と倍増している。

<https://a2a-protocol.org/latest/partners/>

日本からはSoftbankやNTTデータの名前がある。

A2Aプロトコルの当初賛同メンバー(2025年4月時点)

Partners contributing to the Agent2Agent protocol

Google cloudのサイトより

開発者のGoogleに賛同してA2Aプロトコルを推進するパートナー数は増加している



4. A2Aプロトコルによる共同作業

ヘテロジニアスなマルチエージェント

異なるフレームワークで開発の複数エージェント間の通信に

A2Aプロトコルは異なるベンダー等が作成したエージェントを相互接続できるように標準化したことが最大の特徴である。

各社のAIエージェントは独自仕様のAPIやSDKで動作していることが多いため、そのままではエージェント間の通信ができない。

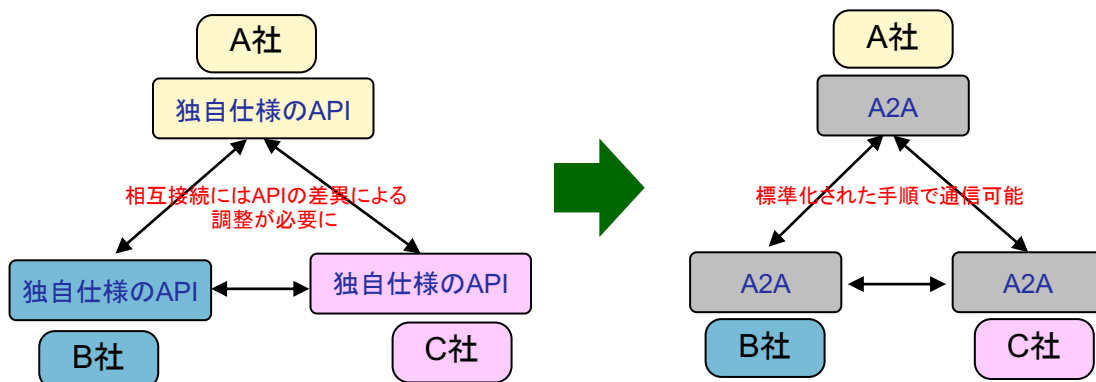
A2Aプロトコルでは、広く利用されているWeb向けのプロトコルであるHTTP/HTTPSと、データマツとしてJSONを用いることでAIエージェント間通信の仕組みを統一している。すなわち、A2AではAIエージェントをHTTPサーバーの形で公開してアクセスできるようにする一方、データ交換形式にはシンプルで軽量な符号化と遠隔呼出し手続きを組合わせたJSON-RPCを採用している。

HTTP: Hyper Text Transfer Protocol HTTPS: HTTP Secure

JSON: JavaScript Object Notation RPC: Remote Procedure Call

A2Aに準拠したエージェントはA2Aで決められた手順によって連携できるので、複数エージェントを組み合わせて便利なサービスが実現できる。

A2Aプロトコルの長所



Agent Cardはエージェントの能力、リンク(エンドポイントURL)、認証要件等をJSON形式で記述した公開ファイル。クライアントはこれを通じて適当な共同作業相手を発見できる。



4. A2Aプロトコルによる共同作業

エージェント・カード

エージェントの素性や能力を記述した公開ファイル

A2Aプロトコルはエージェント間の相互運用を実現するため、次の要素から構成されている。

- ・ A2Aクライアント（クライアント・エージェント）
- ・ A2Aサーバー（リモート・エージェント）
- ・ エージェント・カード
- ・ タスク : IDで識別される作業の基本単位。状態をもちライフサイクルが進行する
- ・ メッセージ : エージェント間通信の都度、user又はagentのロールとパートを持つ
- ・ アーティファクト : タスクの結果で生成される出力(テキスト、画像、データ)でパーツで構成
- ・ パート : メッセージまたはアーティファクト内コンテンツの最小単位

このうちエージェント・カードではエージェントの素性や能力等の基本情報を広報することでエージェント同士が希望する相手エージェントを発見するのを助ける役割を果たす。通常はJSON形式(注)で書かれ、`.well-known/agent.json` のアドレスで公開される。

注) JSON (JavaScript Object Notation) : JavaScriptのオブジェクト記法を基にした軽量で人間が読み書きしやすいテキストベースのフォーマット

Agent Cardの例

```
{
  "name": "経費報告書エージェント",
  "description": "経費報告書の作成と提出を支援します",
  "version": "1.0.0",
  "endpoint": "https://example.com/expense-agent/a2a",
  "capabilities": ["forms", "text"],
  "authentication": {
    "type": "oauth2"
  }
}
```

"name" :Agentの名称
"description" :Agentが実行できる作業概要を紹介
"endpoint" :当該エージェントへのリンク
"capability" :当該エージェントへの出力形式
"authentication" :認証形式

Agent Cardはエージェントの能力、リンク(エンドポイントURL)、認証要件等をJSON形式で記述した公開ファイル。クライアントはこれを通じて適当な共同作業相手を発見できる。



4. A2Aプロトコルによる共同作業

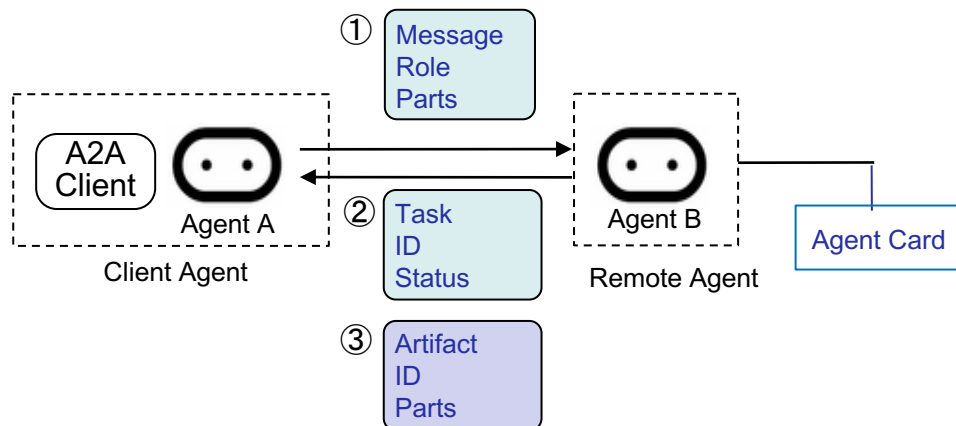
非同期通信モード

複雑な処理の場合は要求側の待合せが不要の通信モードにする

A2Aプロトコルを用いたエージェント間通信は以下のシーケンスで行われる。相手エージェントでは複雑で長時間の処理を要するため、待合せ不要の非同期モードで行うことが多い。

- ・クライアント側エージェント(エージェントA)は、リモート側エージェントが常時広報するデジタル名刺とも言えるエージェントカードを参照して、依頼する作業に適したエージェントBを見つける。
- ・エージェントAはエージェントBのアドレス(URI)宛てに、同期/非同期/ストリーミング/プッシュ通知の別の依頼作業内容(Role, Parts)を記したメッセージを送信する。 :下図①
- ・エージェントBは、これに対して一回のみの返答で済む場合はテキスト/ファイル/JSONデータをエージェントAに返送する。しかし、より複雑な処理や処理時間を要する場合には待合せ不要のタスクオブジェクトとしてエージェントAが希望するタイミングでタスクの進行状況(Status)を問合せできるようにする。 :下図②
このStatusには submitted/working/input-required/completed/failed がある。
- ・タスクが完了した場合には、成果物(Artifact)はParts内のテキスト/ファイル/JSONデータとして示される。 :下図③

非同期通信モードでの通信



非同期通信モードではタスクの進捗状況の問い合わせを適宜実行する。



4. A2Aプロトコルによる共同作業

A2Aプロトコルを使用した協働作業例-1

専門知識/能力を有する外部エージェントとの協働で複雑なタスクを自動化

A2Aプロトコルは次のような形での活用が広がっていくだろう。

1. マルチエージェントでの協働

専門的な知識や能力を有する外部エージェントを活用して複雑なリクエストに応える。

例えば、ユーザがクライアントエージェントに海外旅行の計画作成を依頼する場面では、リクエストに応えるため次のような複数の専門エージェントを利用して調整する。

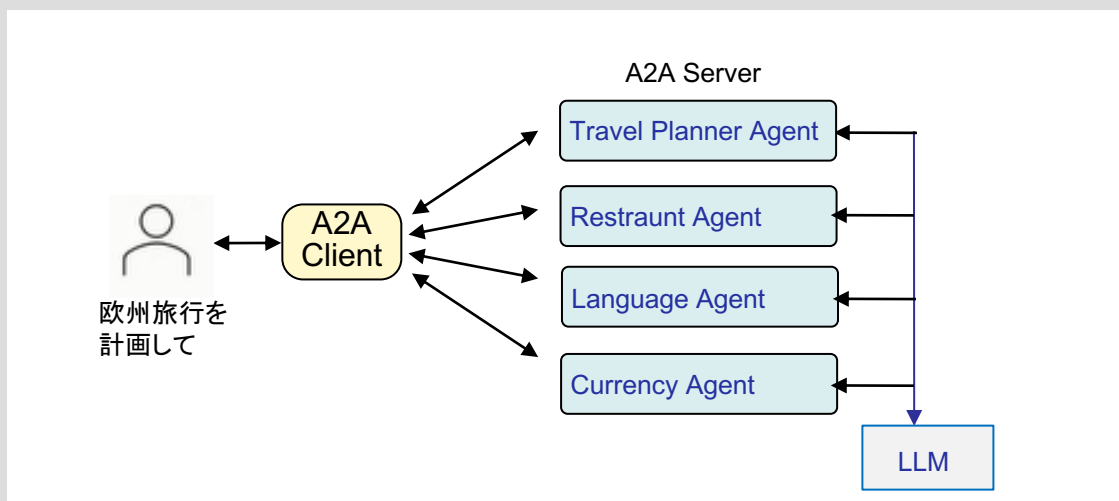
- 航空券予約のエージェント
- ホテル予約のエージェント
- 現地ツアーのおすすめや予約のエージェント

2. マルチモーダルでの対話

顧客サポートの対話サービス等で多様なツールを持つ外部エージェント機能を利用する。

例えば、チャット型のヘルプデスクエージェントがユーザからの質問に対応している間に、その背後ではFAQの検索やITシステムを操作するAIエージェントと連携することで、顧客に最適な回答を提示できるようにする。

旅行計画を作成する例



観光名所・宿泊施設・旅費を提案するTravel Planner Agent, お勧めレストラン情報を提供するRestaurant Agent等の専門のエージェント間が連携して旅行計画を作成する



4. A2Aプロトコルによる共同作業

A2Aプロトコルを使用した協働作業例-2

専門知識/能力を有する外部エージェントとの協働で複雑なタスクを自動化

A2Aプロトコルは次のようなビジネスシーンでの活用が想定される。

<https://feelflow.co.jp/develop/agent2agent/> より

A. カスタマーサポートの自動化

「注文の配送状況を知りたい」との問い合わせに対して:

- 受付エージェント : 顧客の要求を確認
- 注文情報エージェント : 該当する注文を特定
- 発送追跡エージェント : 現在の配送状況を取得
- 回答生成エージェント : 顧客への回答を作成

B. 採用プロセスの効率化

人材採用の効率化のために:

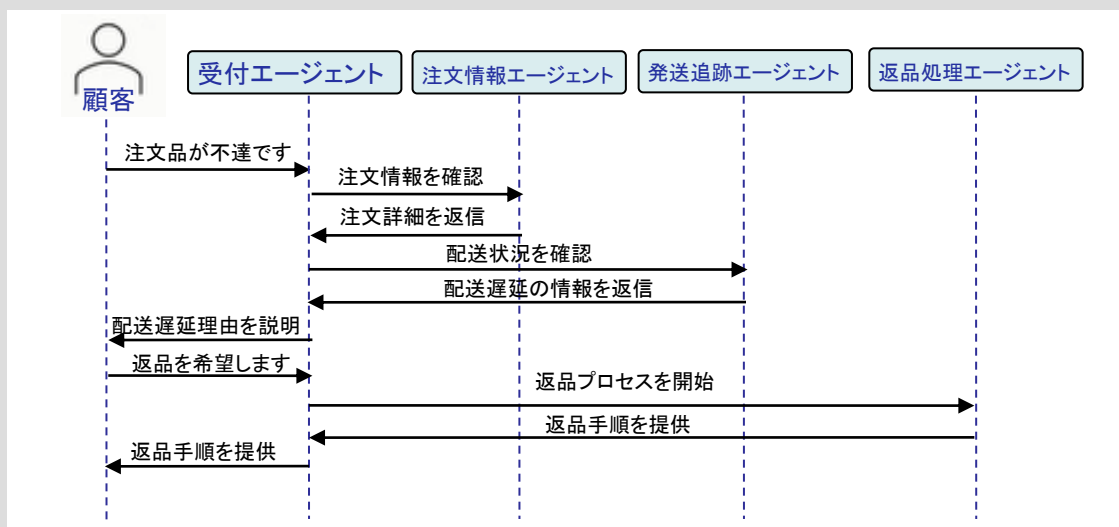
- 要件定義エージェント : 募集条件を整理
- 候補者サーチエージェント: 適切な人材を検索
- 面接予定調整エージェント: 最適な日程を提案
- 評価集約エージェント : 選考結果を報告

C. 財務レポートの自動作成

月次等の財務分析をデータ収集、分析、可視化、報告生成の各エージェント連携で実施

カスタマーサポート自動化の場合

<https://qiita.com/RepKuririn/items/67bde44b58d6512e9409#カスタマーサポート> を元に作成



カスタマーサポートが顧客に対応中に、その背後で注文情報の確認、配送追跡等のエージェントが支援を行う。



【著者略歴】

田代 務

KDDIにて、衛星通信設備の設計や設備導入、研究開発計画の策定、人材育成、海外事業等に従事。ワシントン事務所長、KDDアメリカ副社長、IP事業企画部部長の後にKDDIを退職し、2003年に同僚と株式会社A2A研究所を設立。現在は、主に衛星通信やモバイル通信関係の調査や技術支援を行っている。東大工学部電子工学修士修了。静岡県出身。

主な著書 「どこでもワーク、いつでもラーニング」、「携帯電話の仕組み」、「衛星通信のしくみ」など。
(以上、A2A研究所ウェブサイト www.a2a.jp にて公開中)