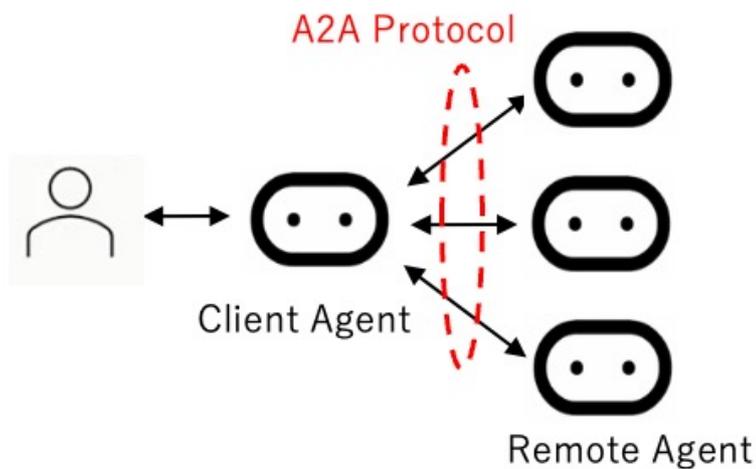


MCP and A2A

~ History of AI toward the automation of collaborative work ~



March 2026

Tsutomu Tashiro, A2A Research

Table of Contents

1 History of AI Development

The 1 st AI boom	3
The era of deep learning	4
Attention Mechanism	5
Businesses driving the spread of AI	6

2 Generative AI

LLM (Large-scale Language Models)	7
Generative AI Service Models	8
Open-source generative AI service	9
Image Generative AI	10

3 AI-powered Streamlining of intellectual tasks

AI Agent	11
MCP (Model Context Protocol)	12
Streaming by MCP	13
A2A (Agent to Agent Protocol)	14
Components using A2A Protocol	15
Supporters of the A2A Protocol	16

4 Collaborative work by A2A Protocol

Heterogeneous Multiagent	17
Agent Card	18
Asynchronous Communication Mode	19
Examples of Collaborative work using A2A -1	20
Examples of Collaborative work using A2A -2	21



1. History of AI Development

The 1st AI boom

The first AI boom began in 1960s.

The term "artificial intelligence" (AI) originated in a 1956 workshop held at Dartmouth College, where Professor John McCarthy proposed, "Since all learning and intelligence functions can be accurately simulated by machines, I would like to call this artificial intelligence."

■ The First AI Boom: "The Age of Inference and Search"

From the 1950s to the 1970s, even the most powerful computers were limited to very specific problems, such as proving geometric theorems or learning English conversation. However, the intellectual capabilities of these computers were considered astounding, so optimistic predictions about future advances in computing power spread among AI researchers.

In Japan, Shunichi Amari first successfully developed the "stochastic gradient descent method for multilayer perceptron," which became the foundation of modern AI in 1967. Hinton and others subsequently rediscovered it as the backpropagation algorithm.

■ The Second AI Boom: "The Age of Knowledge Engineering"

In the 1980s, expert systems, a type of AI, were introduced in many companies. These systems answered questions and solved problems in specific areas, and although there were some successful examples, such as identifying compounds from spectroscopic measurement results and diagnosing blood disorders, they did not see widespread application.

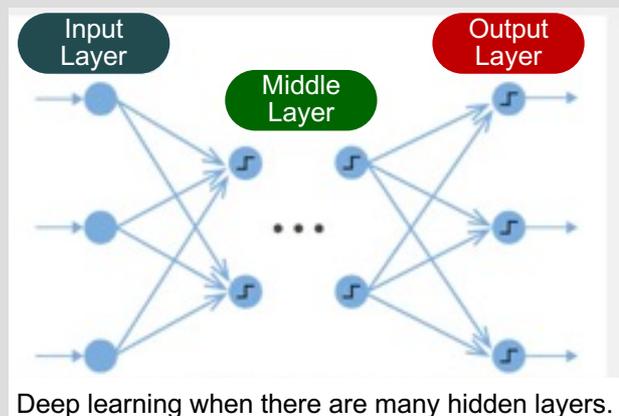
In 1981, the MITI launched the Fifth Generation Computer Project, aiming to achieve human-like reasoning through natural language dialogue with humans, machine translation, and image recognition, but were unsuccessful.

Multilayer Perceptron

◆ Backpropagation is a learning method that adjusts the network weights W by working backward from the output layer to minimize the difference (error) between the output value and the correct answer.

◆ Stochastic gradient descent is a specific method of backpropagation. It finds the direction (gradient) where the error is smallest and adjusts the network weights w in that direction. By randomly selecting a portion of the training data, randomness can be introduced in the gradient direction to search for the optimal solution.

【 Image of a perceptron intermediate layer 】



The multilayer perceptron, which can be considered the prototype of the core of current AI, was conceived half a century ago.



1. History of AI Development

The era of deep learning

The 2024 Nobel Prize in Physics and Chemistry were both AI-related

■ The Third AI Boom: "The Age of Deep Learning"

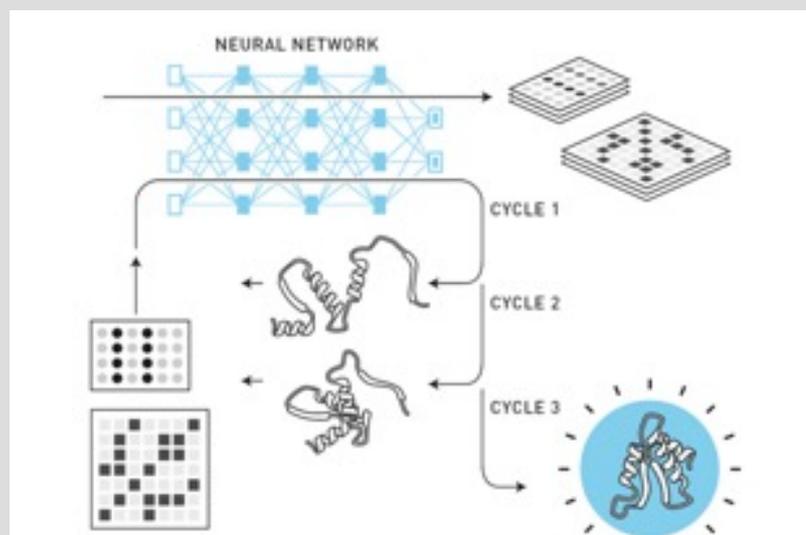
In 2006, Geoffrey Hinton invented deep learning, which allows for the extraction of features without human intervention, marking a major breakthrough. In 2012, Google announced its success in recognizing cats from YouTube images using deep learning, attracting widespread attention.

■ The Fourth AI Boom: "LLM and Generative AI"

In 2020, OpenAI developed the automatic language processing program GPT-3. It used a transformer based on neural networks and a self-attention mechanism, and a large-scale language model (LLM). Because the learning and inference of this LLM requires high-speed parallel calculations of a huge number of floating-point numbers, supercomputers were initially attempted. Later, when it was discovered that it could be handled by NVIDIA's graphics processing unit (GPU) for game consoles, its use among researchers rapidly spread.

In 2024, Geoffrey Hinton and others received the Nobel Prize in Physics for "fundamental discoveries and inventions that make machine learning using neural networks possible." In the same year, the CEO of Google and researchers received the Nobel Prize in Chemistry for "protein structure prediction using AI," giving a strong impression on the world that generative AI has become an important foundation supporting modern society. On the other hand, the dark side of this is also becoming stronger, such as the misuse of generative AI to spread false information.

Illustrated diagram of a method for predicting protein structure from amino acid sequences.



(Provided by the Royal Swedish Academy of Sciences/Nobel Foundation)

The 2024 Nobel Prizes in Physics and Chemistry were both awarded for their connection to the practical application of AI.



Attention Is All You Need

One of the most successful early applications of AI technology was machine translation. A 2017 paper by Google researchers on Transformer, a deep learning model that achieved high-performance machine translation, expressed its core principle as follows: "Attention Is All You Need."

This Transformer has the following general structure:

- It divides the source text into tokens (roughly equivalent to words) and converts them into token sequences (vectors).
- It uses recurrent or convolutional neural networks (NNs) on both the source and target text sides.
- An attention mechanism is applied to both NNs.

The attention mechanism is a function that draws attention to important tokens in a sequence of tokens extracted from a text by quantifying the degree of relevance (or similarity) between neighboring tokens on a scale of 0 to 1.

Because this model is parallelizable, using a GPU suitable for parallel processing significantly reduced training time, allowing us to achieve high machine translation quality (BLEU: Bilingual Evaluation Understudy score) in a short amount of time.

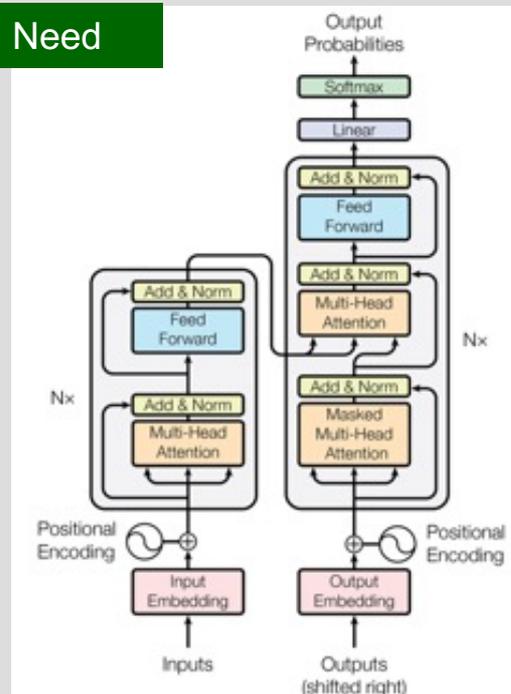
Attention Is All You Need

Figure 1 from the landmark paper "Attention Is All You Need." by Google researchers: The Transformer – model architecture.

The left side of the diagram shows the decoder for the source text, and the right side shows the encoder for the translated text.

Within the neural network (Nx), the relationships between token sequences in a text are represented as matrices, and the degree of relationship is determined by calculating the determinant, resulting in a value called Attention, which is then converted into a matrix.

For example, in the sentence "Please summarize the text," we understand that the tokens "text" and "summary" are important.



Attention, the act of focusing on specific parts of an image or text, is something humans do every day.



1. History of AI Development

Businesses driving the spread of AI

OpenAI and Anthropic

In recent years, news reports on AI trends have dominated the media. For example, it's been claimed that AI's software coding capabilities have surpassed those of skilled engineers, meaning engineers no longer code but primarily check the AI.

A few years ago, the so-called singularity (technological singularity), where AI capabilities surpass human capabilities, was predicted to occur in the mid-2040s. However, recent predictions suggest it will happen much earlier, around the early 2030s.

Leading the way in these disruptive AI innovations are OpenAI and Anthropic.

OpenAI quickly spread its chat application, GPT, which can instantly and almost accurately answer a wide range of human questions, throughout the world.

Since then, many businesses have emerged offering generative AI services suitable for generating text, images, videos, audio, music, and software programs, having serious impact on a wide range of intellectual workers, from teachers and consultant, to creators.

On the other hand, Anthropic, which specializes in automating various business processes for enterprises, is seen as a formidable competitor to SaaS (Software as a Service) providers. The emergence of AI agent protocols for inter-company business, coupled with this development, may bring about significant innovation in both intra and inter-company business. In February 2026, the "Anthropic shock" occurred, where the stock prices of companies developing related businesses plummeted due to concerns that high-performance AI would seize demand for traditional services.

Open AI and Anthropic

	Open AI: Sam Altman	Anthropic: Dario Amodei
	<p>1985 born in St. Louis. His mother was a dermatologist of Jewish descent.</p> <p>~2005 studied computer science at Stanford University.</p> <p>2015 co-founded the research company OpenAI with Elon Musk and Peter Deal, and became the CEO.</p> <p>2023 left OpenAI and announced he would be joining Microsoft.</p> <p>2024 later returned to OpenAI. OpenAI operates as both the non-profit OpenAI Foundation and the for-profit Open Group.</p>	<p></p> <p>earned a Bachelor of Science degree in Physics from the Univ of California and Stanford Univ, and a PhD in Physics from Princeton Univ for his research on neural circuits.</p> <p>2014 after joining Baidu, moved to Google.</p> <p>2016 joined OpenAI.</p> <p>2021 co-founded Anthropic with his sister Daniela and other former members of OpenAI.</p> <p>2023 The U.S. Senate Judiciary Committee warned of the dangers AI poses to weapons development and control.</p>

After leaving OpenAI, Anthropic will focus on business innovation within and between companies.



2. Generative AI

LLM (Large-scale Language Models)

Neural networks with hundreds of hidden layers are common.

A Large Language Model (LLM) is an AI model that learns from massive amounts of text data using deep learning techniques to achieve a high level of language understanding and generation. The most common model is a multi-layer Neural Network (NN) with hundreds of hidden layers between the input and output layers.

The numerical values of the hundreds of millions to tens of billions of parameters in a multi-layer NN are determined through learning from large amounts of data (articles on the web, books, news, conversation records, etc.), and the LLM will have the following functions.

- Answering questions
- Summarizing texts
- Machine translation
- Predicting the next prompt (input)
- Keyword extraction
- Program bug checking, etc.

While LLM is a model specifically designed for text understanding and generation, it is also applied to methods for automatically generating non-text content such as images, videos, audio, and is often collectively referred to as generative AI.

Main LLM				
Models	Developer	Year	Parameters	Features
BERT	Google	2018	340 mil	Initial model with improved accuracy through increased dataset size
GPT-4	OpenAI	2023	private	Trained GPT-3 with non-text data as images and audio
PaLM	Google	2022	540 bn	Much increased number of Transformer parameters
LLaMA	Meta	2023	7~65 bn	Achieves performance equivalent to GPT-3 with fewer parameters, open-sourced and released on GitHub
Claude	Anthropic	2023	private	Model developed by GPT-2 and GPT-3 developers
Gemini1.5	Google	2024	private	Multimodal model developed by Google DeepMind
OpenELM	Apple	2024	1 bn	Creates multiple layers and efficiently places parameters within each layer
GPT-5	OpenAI	2025	private	Usable in fields such as coding, mathematics, and document creation
Tsuzumi-2	NTT	2025	0.6~7 bn	Lightweight and supports Japanese and English
cotomi	NEC	2024	13 bn	Lightweight and supports Japanese
ELYZA	Tokyo Un	2025	70 bn	Japanese-focused model developed by an AI company originating from the Matsuo Laboratory*

*) KDDI subsidiary

Currently, mainstream LLMs have hundreds to hundreds of billions of parameters.



2. Generative AI

Generative AI Service Model

From text generation to video and music generation

Generative AI, a technology that automatically creates text, images, music, videos, and other content from training data, is rapidly evolving.

For example, the latest version of OpenAI's chat application, GPT-5, enables the generation of images and audio in addition to text, as well as advanced reasoning.

By prompting for appropriate context and specific requests, it can even create visually appealing websites and games.

The industry is truly a veritable explosion of talent, with companies providing AI generation services including AI-focused ventures like OpenAI and Anthropic, big tech giants like Google (whose parent company is Alphabet) and Microsoft, and design and software giant Adobe, as well as Stable Diffusion, which provides open-source AI specializing in image generation, and VOICEVOX, which also generates music using open-source software.

The following uses the four types of generated products: text, images, videos, and audio/music:

- * Text generation: Creation of reports and emails, summarization of documents, translation, and creation of meeting minutes.
- * Image generation: Creation of images for social media posts and advertisements.
- * Video generation: Creation of promotional videos and educational content.
- * Audio/music generation: Creation of narrations and audio guides, and music production.

Generative AI Services				Created based on information from the Fujifilm website
Product	Service	Developer	Monthly charge	Feature
sentence	ChatGPT	OpenAI	Plus:20 \$, Pro:200 \$	Highly versatile, supports images and audio, GPT-5 for web, apps, and games
	Gemini	Google	Pro:2900 Yen	Excellent for searching the latest information and processing long texts
	Claude	Anthropic	Pro:17 \$, Max: 100 \$	Ethically accurate, natural Japanese
	Copilot	Microsoft	Personal:2130 Yen Premium:3200 Yen	Integrates with Office tools
image	Stable Diffusion	Open Source	Pro:7 \$, Max: 14 \$	Models can be freely adjusted on-premise
	Adobe Firefly	Adobe	Standard: 1580 Yen Pro:4780 Yen	Integrates with Adobe tools, prioritizing security
video	Sora	OpenAI	ChatGPT Plus needed	High prompt understanding for visual expression
	Runway		Standard:15 \$, Pro:35 \$	Video generation from images
Voice-Music	VOICEVOX	Open Source	Free	Japan-original speech synthesis software
	Suno AI		Pro:1500 Yen	Generates music from prompts in minutes

Generative AI services capable of efficiently generating all forms of intellectual output are emerging one after another.



2. Generative AI

Open-source generative AI service

From text generation to video and music generation

Open-source AI is AI whose source code is publicly available, allowing anyone to freely use, modify, and distribute it.

Software developer platforms such as GitHub make AI algorithms, pre-trained models, and datasets publicly available. Using libraries like the following, volunteer enthusiast communities are developing practical AI solutions based on existing works.

- TensorFlow: Computation for mathematical formulas and neural network (NN) processing
- PyTorch: Python-based machine learning library with intuitive operation
- Scikit-learn: Python library for machine learning and predictive data analysis
- OpenCV: Provides various algorithms for real-time image and video processing

Compared to proprietary AI models (closed source) developed by companies, these offer greater customization flexibility, can be run in a local environment, resulting in lower costs and better privacy protection. However, depending on the licensing model, issues of liability and sustainability may arise, and high-performance GPUs may be required.

For general use only, there are services that generate images and music, such as:

- Stable Diffusion: image generation AI
- VOICEBOX : singing voice synthesis AI

Examples of open-source generative AI services



Top: Response to the "Cat wearing sunglasses" prompt in Stable Diffusion (GRUSOROBAN website) <https://soroban.highreso.jp/article/article-034>
Right: VOICEVOX - Free text-to-speech by Ryusei Aoyama VOICEVOX website <https://voicevox.hiroshiba.jp>



Even beginners can generate images and music easily using open-source AI services.



2. Generative AI

Image Generation AI

The diffusion model is a typical approach.

Image generation AI works by pre-training a vast number of image-text pairs (hundreds of millions to billions) using deep learning. It then searches for features based on the pre-trained data from user-input text and generates images from matching features.

For example, inputting "cat, napping" would generate photos or illustrations of a napping cat. Various services offer the option to select the type of generated image, such as live-action photos, animations, or 3D models.

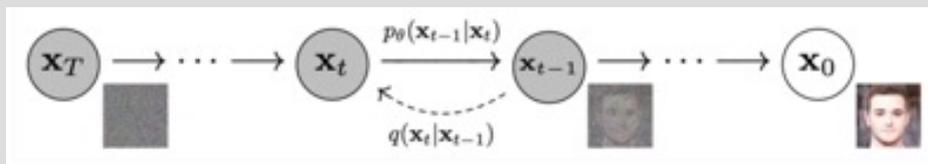
There appear to be various methods for this image generation AI.

The most prominent is the Diffusion Model, which is used by Stable Diffusion, an open-source image data generation service.

In the first step of the Diffusion Model, noise is gradually added to real image data similar to the requested image to create a noisy image. Next, during the process of gradually removing noise from this noisy image, the AI uses a large amount of previously learned images and prompt instructions to fill in the shapes and colors, resulting in a realistic and natural requested image.

While AI-generated images allow anyone to easily create high-precision images, the lack of established legal frameworks means that misuse could infringe on copyrights and intellectual property rights. For example, this could occur when generating and using images that closely resemble existing characters or illustrations.

Image generation using diffusion models



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



For example, generate the requested image on the right based on the similar image on the left.

The requested image is generated by adding noise to a similar real image, then gradually correcting its shape and color based on prompt instructions during removing that noise.



3. AI-powered Streamlining of Intellectual tasks

AI Agent

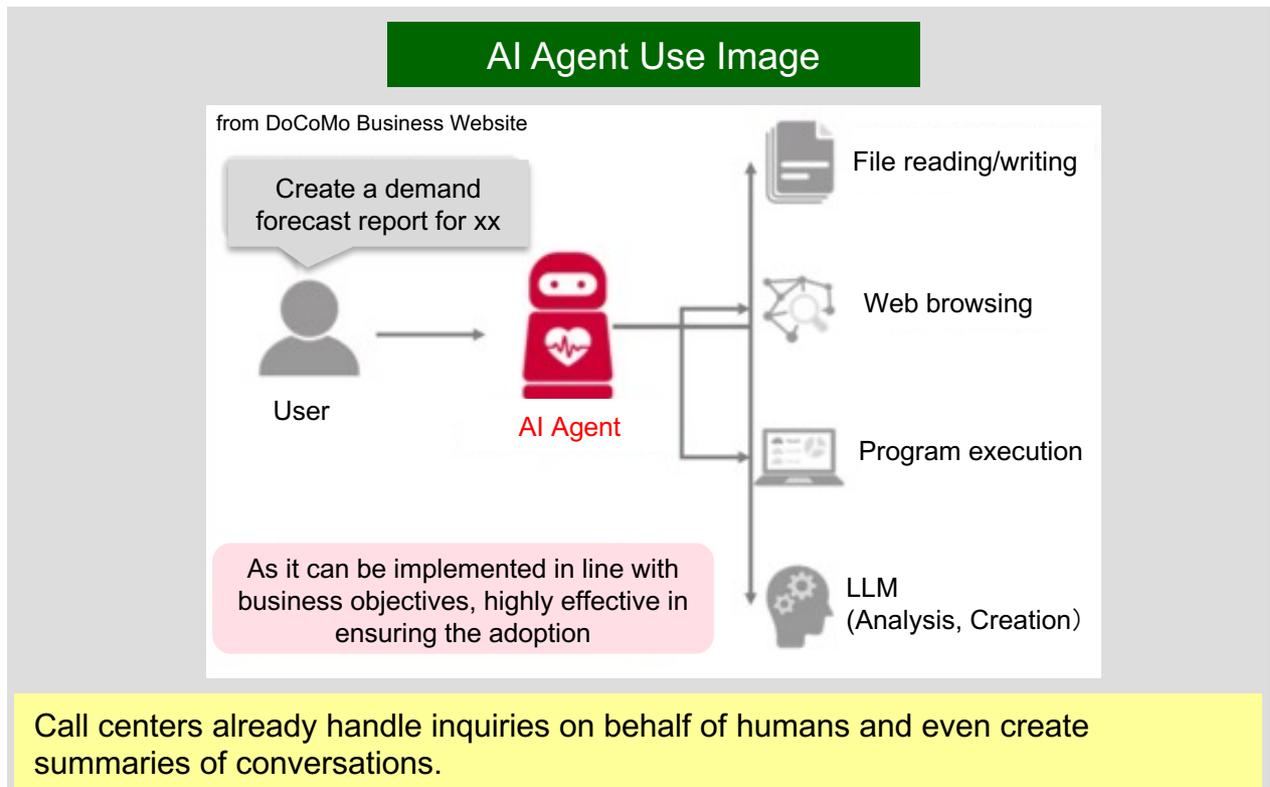
Software that automates tasks to replace human work

AI agents are autonomous software that plan their own work objectives, search for necessary information, and complete tasks using various tools without requiring detailed human instructions.

For example, an AI agent in a call center could automatically perform the following tasks in place of a human:

- Search FAQs in real time based on the content of customer conversations
- Summarize conversations after customer calls and create a knowledge base for internal sharing as needed

Within companies, in addition to handling inquiries, on-site tasks such as accounting and maintenance, as well as liaison work between multiple departments, can easily be replaced by AI agents. As shown in the diagram below, AI agents in internal operations automatically complete tasks on behalf of users by appropriately utilizing multiple tools such as web searches, file input/output, and program execution.



Call centers already handle inquiries on behalf of humans and even create summaries of conversations.



3. AI-powered Streamlining of Intellectual tasks

MCP (Model Context Protocol)

Anthropic announced it.

MCP (Model Context Protocol) is a standard protocol for connecting large-scale language models (LLMs) with external systems, announced by Anthropic in November 2024.

For business-oriented AI applications such as AI agents with built-in LLMs, a mechanism for appropriately supplying necessary context* from external tools is crucial.

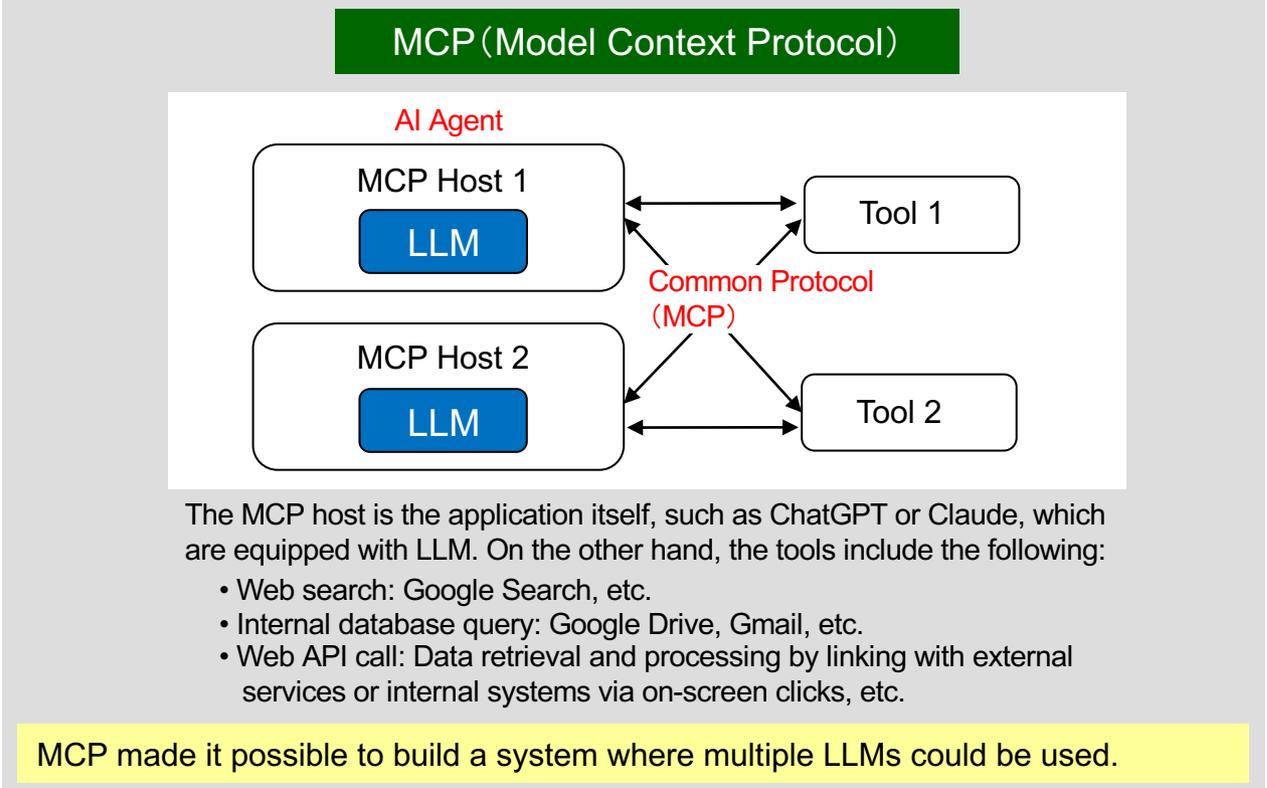
MCP (Model Context Protocol) is a protocol that standardizes this context supply.

*) Context: Supplementary information necessary for responses by the LLM. It includes the conversation history preceding the conversation, the background and purpose of instructions, information such as time, place, organization, and the user's past behavior and preferences.

Until then, methods developed by OpenAI, Google, and others required different implementations due to their proprietary specifications.

MCP has made it possible to build systems that can utilize multiple LLMs. Since different "devices" (LLMs) can be connected with the same "cable" (MCP), it is sometimes likened to an "AI USB-C port."

In addition to the protocol, Anthropic also announced a software development kit (SDK) that supports languages such as Python. This enables collaboration between diverse AI applications and data sources, thus evolving AI from a mere conversational application into an intelligent work partner.



MCP made it possible to build a system where multiple LLMs could be used.



3. AI-powered Streamlining of Intellectual tasks

Streaming by MCP

Large sized data is divided into smaller parts and transmitted multiple times.

MCP (Model Text Protocol) incorporates streaming capabilities, familiar from video and music, in addition to HTTP (Hypertext Transfer Protocol), a request/response type communication protocol between web browsers and servers.

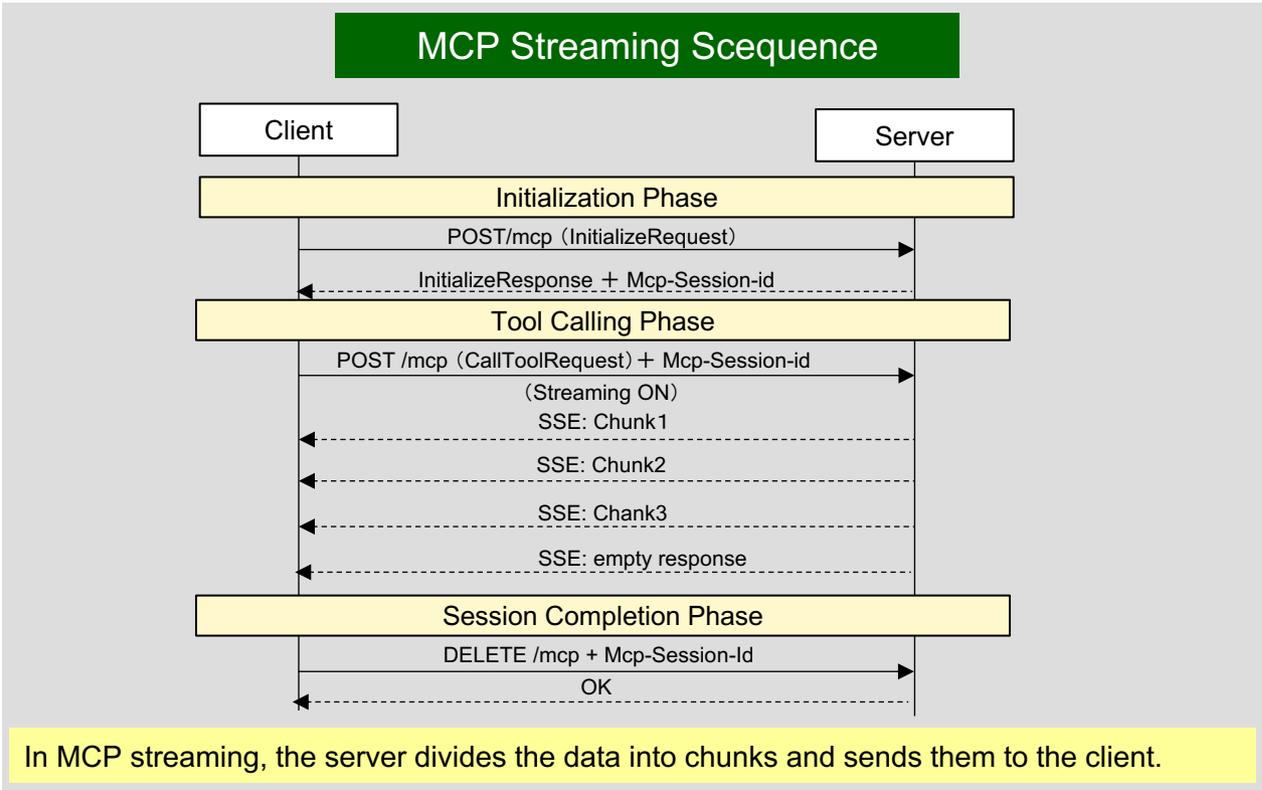
While the HTTP request-response model was often sufficient, it was inconvenient when real-time processing was required or when large amounts of data processing would cause long wait times for the client.

Therefore, MCP includes streaming capabilities as a standard feature and further enables the creation of interactive applications by adding bidirectional communication such as notifications of processing progress and distribution of intermediate results.

The diagram below shows an example of a streaming sequence in MCP, where the server divides the data into small chunks and sequentially sends each chunk to the client as an SSE (Server-Sent Event).

The final SSE signals "transmission complete."

Many generative AI servers maintain a connection with the client and push data little by little to provide real-time response notifications.





3. AI-powered Streamlining of Intellectual tasks

A2A (Agent to Agent Protocol)

Developed by Google in 2025

A world is dawning where AI agents from multiple companies and organizations collaborate to perform complex tasks.

A2A (Agent to Agent Protocol) was developed by Google in 2025 as a protocol for AI agent communication and subsequently donated to the Linux Foundation.

The A2A protocol has the following features:

- Simple design based on HTTP and JSON-RPC
- Supports real-time communication via streaming
- Asynchronous task management via push notifications
- Robust security features for enterprises

The A2A protocol also has the following components:

- Client Agent: Requests work from another agent on behalf of a user (person)
- Remote Agent: Processes the requested work by publishing an HTTP endpoint, etc.
- Agent Card: Identity document showing the agent's capabilities, authentication method, etc.

In the A2A protocol, when a request is sent to the other agent, a Message or Task is sent back as a response. Message is a response that is completed in a single exchange. On the other hand, Task has a state, and the progress of the process can be managed by changing the state from submitted → working → completed.

Comparison of Message and Task in the A2A protocol

Judgement Item	Message	Task
Processing Time	Short duration (seconds)	Long duration (minutes ~ hours)
Status Management	Not required	Required
Progress Check	Not required	Possible
Cancellation	Not possible	Possible
Deliverables Management	Response only	Save as Artifact
Resource Usage	Light	Heavy

- Message example**
- Greeting with "Hello"
 - Answering simple questions
 - Notification of calculation results
 - Checking status

- Task example**
- Creating reports and documents
 - Analyzing large amounts of data
 - Creating images and videos
 - Processing multiple steps

Messages are like question-and-answer responses, while Tasks are for tasks that require complex processing or state management.



3. AI-powered Streamlining of Intellectual tasks

Components using the A2A protocol

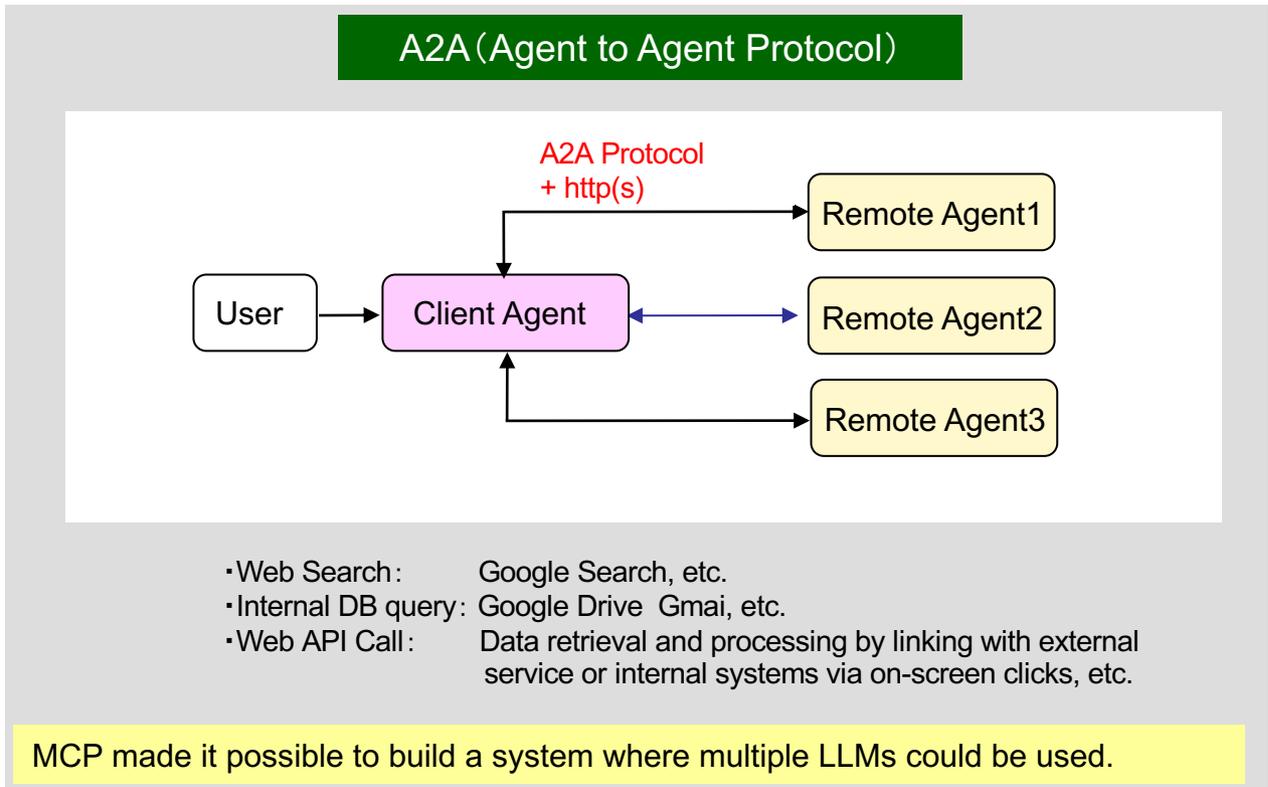
Adopts a client-server model

A2A (Agent to Agent Protocol) is a protocol used to automatically perform collaborative work by requesting one or more external agents to execute tasks on the client side. While MCP is primarily used within enterprises when AI applications with built-in LLMs utilize external services and resources such as data sources and tools, A2A is used for collaborative work with external AI agents.

One use case is when a user creates a travel plan.

A travel agency, upon receiving this request from a client, would then commission various agents specializing in specific areas such as hotels and other accommodations, and transportation like trains and airplanes, to create the overall itinerary and expense report.

Data exchange typically uses HTTP (Hypertext Transfer Protocol), a communication protocol for transferring files and images between a web browser and a server, or HTTPS (HTTP Secure), which encrypts that communication.





3. AI-powered Streamlining of Intellectual tasks

Supporters of the A2A Protocol

The number is increasing

The AI agent-to-agent protocol A2A, developed by Google in 2025 with their expertise, complements MCP, which was originally developed by Anthropic, and has garnered widespread support from the outset.

A2A is designed to address the challenges of deploying large-scale multi-agent systems, and its interoperability, which standardizes agent management across diverse platforms and cloud environments, seems to be particularly noteworthy.

Google stated the following regarding the release of the A2A protocol.

"A2A provides developers with the ability to build agents that can connect with other agents that comply with the protocol, giving users the flexibility to combine agents from diverse providers."

The initial members of the A2A protocol, announced on the Google Cloud website in April 2025, were approximately 60 US-based companies. However, according to the following site, as of March 2026, the number of participants from other countries had increased, doubling to a total of approximately 120 companies.

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

From Japan, there are names like Softbank and NTT Data.

Original supporters of the A2A protocol (as of April 2025)



from Google cloud site

The number of partners supporting the A2A protocol is increasing.



4. Collaborative work by A2A Protocol

Heterogeneous multiagent

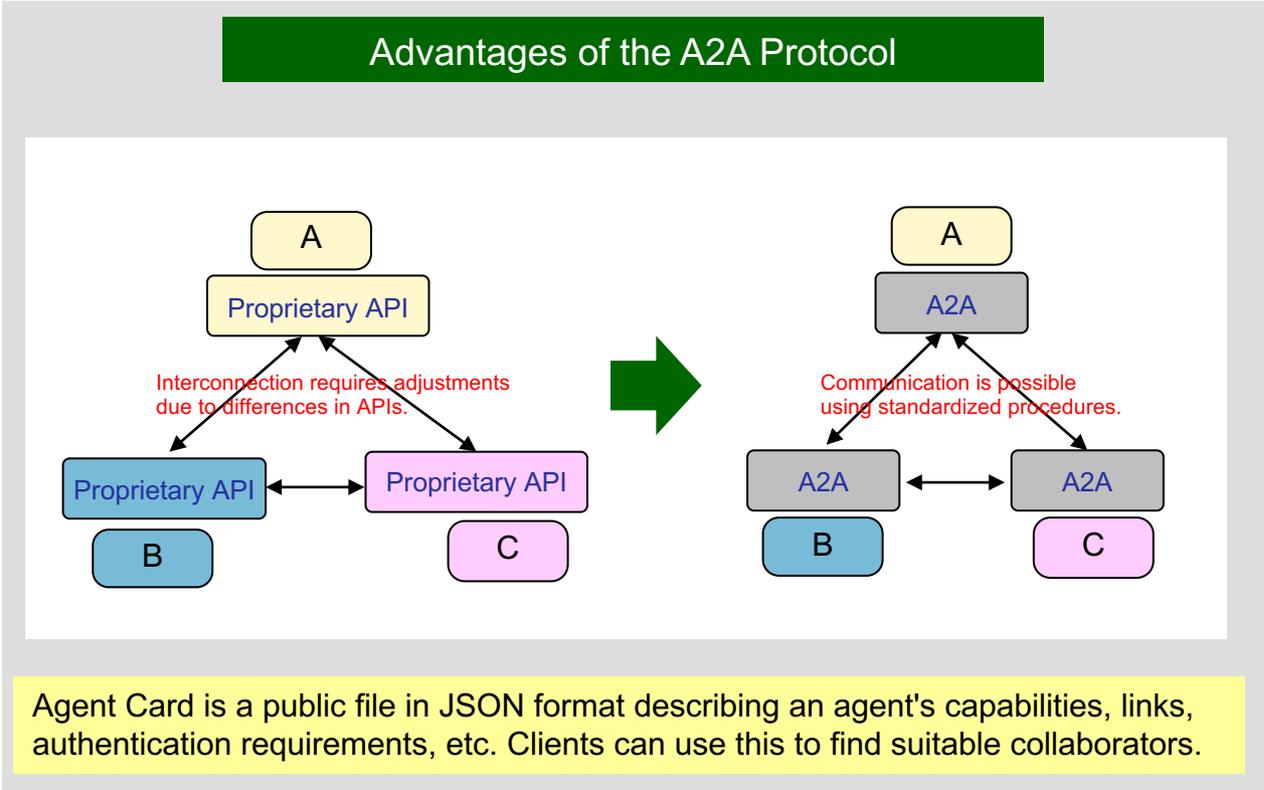
Communicate between multiple agents developed in different frameworks

The most significant feature of the A2A protocol is that it standardizes agents created by different vendors, enabling them to interconnect. Because each company's AI agents often operate using proprietary APIs and SDKs, they cannot communicate with each other directly.

The A2A protocol unifies the mechanism for communication between AI agents by using HTTP/HTTPS, a widely used web protocol, and JSON as the data format. In other words, A2A exposes AI agents as HTTP servers for access, while employing JSON-RPC, a simple and lightweight encoding and remote calling procedure, as the data exchange format.

HTTP: Hyper Text Transfer Protocol HTTPS: HTTP Secure
JSON: JavaScript Object Notation RPC: Remote Procedure Call

Agents compliant with A2A can collaborate according to procedures defined by A2A, allowing for the combination of multiple agents to create convenient services.





4. Collaborative work by A2A Protocol

Agent Card

Public files describing the agent's background and abilities

The A2A protocol, designed to enable interoperability between agents, consists of the following elements:

- A2A Client (Client Agent)
- A2A Server (Remote Agent)
- Agent Card
- Task: A basic unit of work identified by an ID. It has a state and a lifecycle.
- Message: Has a user or agent role and part for each agent-to-agent communication.
- Artifact: Output (text, image, data) generated as a result of a task, composed of parts.
- Part: The smallest unit of content within a message or artifact.

Agent cards, in particular, play a role in helping agents find each other by publicizing basic information such as their background and abilities.

These are usually written in JSON format (Note) and published at the address of:
.well-known/agent.json.

Note) JSON (JavaScript Object Notation) is a lightweight, human-readable, text-based format based on JavaScript object notation.

Agent Card Example

```
{
  "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 Name
"description" : Outlines the tasks that the agent can perform.
"endpoint" : Link to the Agent
"capability" :Output format to the agent
"authentication" :Authentication format

An Agent Card is a public file in JSON format that describes an agent's capabilities, links (endpoint URLs), authentication requirements, etc. Clients can use this to find suitable collaborators.



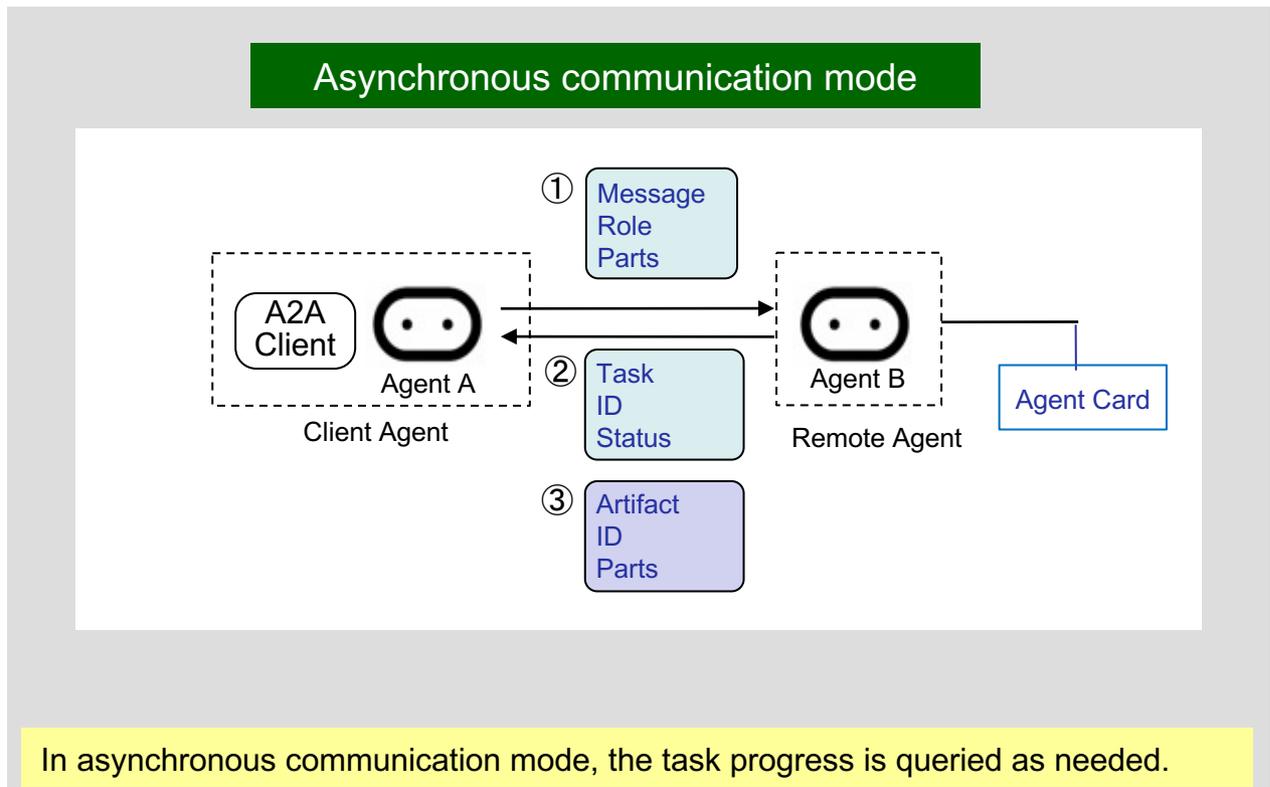
4. Collaborative work by A2A Protocol

Asynchronous communication mode

For complex processing, a mode without waiting is used.

Agent-to-agent communication using the A2A protocol follows the sequence below. Because the remote agent requires complex and time-consuming processing, it is often performed in asynchronous mode, eliminating the need for waiting.

- The client agent (Agent A) refers to the agent card - a digital business card so to speak - which the remote agent constantly advertises, to find Agent B suitable for the requested task.
- Agent A sends a message to Agent B's address (URI) containing details of the requested task (Role, Parts) via synchronous/asynchronous/streaming/push notification. (See Figure ① below)
- If Agent B only needs to respond once, it will send text/file/JSON data back to Agent A. However, if the processing is more complex or requires more processing time, it will be treated as a task object that does not require waiting, allowing Agent A to query the task's progress (Status) at the time of their choice. (See Figure ② below.) Status can be submitted/working/input-required/completed/failed.
- When a task is completed, the artifact will be shown as text/file/JSON data within Parts. (See Figure ③ below)



In asynchronous communication mode, the task progress is queried as needed.



4. Collaborative work by A2A Protocol

Examples of collaborative work using the A2A protocol -1

Automate complex tasks by collaborating with external agents

The A2A protocol is expected to see widespread use in the following ways:

1. Multi-agent collaboration: Utilizing external agents with specialized knowledge and capabilities to respond to complex requests.

For example, when a user requests a client agent to plan an overseas trip, multiple specialized agents will be used and coordinated to fulfill the request:

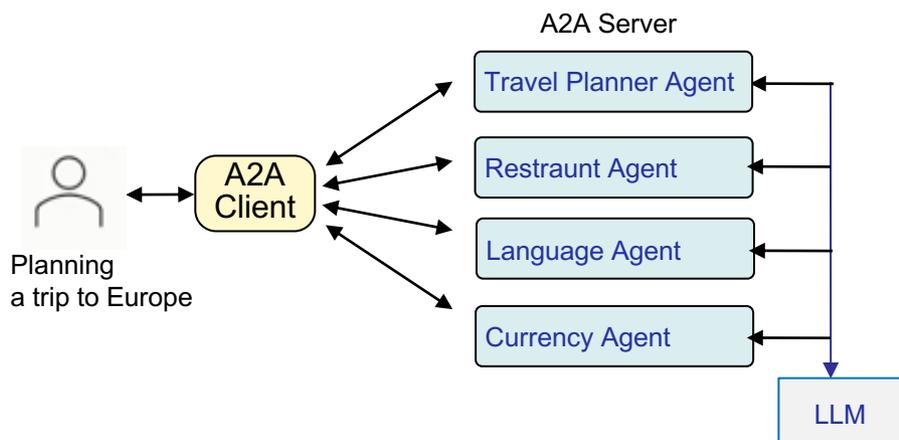
- Flight booking agent
- Hotel booking agent
- Agent for recommending and booking local tours

2. Multimodal Dialogue

Utilize external agent functions with diverse tools in customer support dialogue services, etc.

For example, while a chat-type help desk agent is responding to a user's question, an AI agent that searches FAQs and operates IT systems can be integrated in the background to provide the customer with the most appropriate answer.

Example of Creating a Travel Plan



Travel planning is created through collaboration between specialized agents such as Travel Planner Agents who suggest tourist attractions, accommodations, travel expenses, and restaurants.



4. Collaborative work by A2A Protocol

Examples of collaborative work using the A2A protocol -2

Automate complex tasks by collaborating with external agents

The A2A protocol is expected to be used in the following business scenarios:

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

A. Automated Customer Support

In response to an inquiry such as "I want to know the delivery status of my order":

- Reception Agent: Confirms the customer's request
- Order Information Agent: Identifies the relevant order
- Shipping Tracking Agent: Retrieves the current delivery status
- Response Generation Agent: Creates a response for the customer

B. Streamlining the Recruitment Process To streamline talent acquisition:

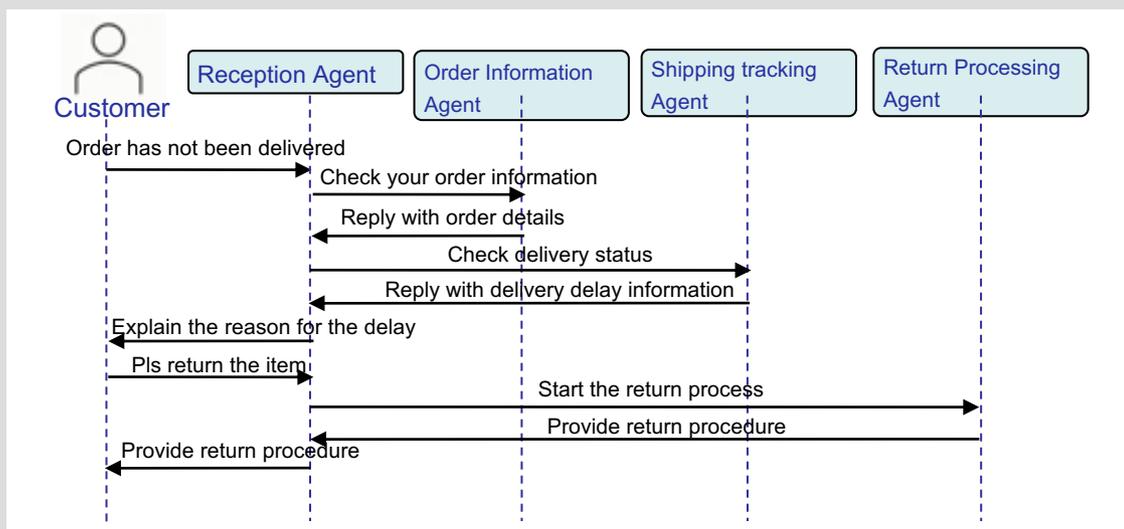
- Requirements Definition Agent: Organize recruitment conditions
- Candidate Search Agent: Search for suitable candidates
- Interview Scheduling Agent: Propose optimal dates
- Evaluation Aggregation Agent: Report selection results

C. Automated creation of financial reports:

Monthly and other financial analyses are performed through the integration of agents for data collection, analysis, visualization, and report generation.

In the case of customer support automation

Based on <https://qiita.com/RepKuririn/items/67bde44b58d6512e9409#カスタマーサポート>



While support staff assisting customers, agents provide support in the background by verifying order information, tracking deliveries, and performing other tasks.