

生成 AI を活用するコミュニケーションツールの 実現に向けた考察と試み

Considerations and trials toward the implementation of communication tools that utilizes generative AI

中口 孝雄 (京都情報大学院大学, 未来環境ラボ)

秋山 功 (BIPROGY 株式会社, 未来環境ラボ)

三浦 仁 (BIPROGY 株式会社, 未来環境ラボ)

前納 一希 (京都コンピュータ学院, 未来環境ラボ)

橋本 昇 (京都情報大学院大学, 未来環境ラボ)

江尻 秀彰 (BIPROGY 株式会社, 未来環境ラボ)

Takao Nakaguchi (The Kyoto College of Graduate Studies for Informatics., Future Lab.)

Isao Akiyama (BIPROGY Inc., Future Lab.)

Hitoshi Miura (BIPROGY Inc., Future Lab.)

Kazuki Maeno (Kyoto Computer Gakuin., Future Lab.)

Noboru Hashimoto (The Kyoto College of Graduate Studies for Informatics., Future Lab.)

Hideaki Ejiri (BIPROGY Inc., Future Lab.)

Abstract

新型コロナウイルス感染症の流行に伴ってコミュニケーションツールの導入が進み、新たなツールも多く登場した。その後ウィズコロナ時代に入り、コミュニケーションツールを巡る動向は落ち着いたように思われる。一方で画像生成や ChatGPT に代表されるテキスト生成などの生成 AI 技術が急速に発展し、生産性や品質の向上が期待されるとともに、雇用の喪失や悪用などマイナス面にも関心が高まっている。様々なソフトウェアで生成 AI の採用や活用が進んでおり、コミュニケーションツールもその一つとなっている。未来環境ラボでは、オンラインコミュニケーションツールのあるべき姿や活用方法に対する議論とプロトタイピングを重ねてきた。生成 AI の採用についても議論やコミュニケーションツールへの導入の試みを行っており、本稿ではその一端を紹介する。

With the outbreak of COVID-19, communication tools have become popular, and many new tools have appeared. Then, with the advent of the with-coronavirus era, the trends surrounding communication tools have settled down. On the other hand, the rapid development of generative AI technologies, such as image generation and text generation represented by ChatGPT, is expected to improve productivity and quality. At the same time, there is growing interest in the negative aspects, such as job loss and misuse. Various software applications are adopting and utilizing generative AI, and communication tools are one of them. At the Future Lab, we have been discussing and prototyping what online communication tools should be and how to use them. We have also been discussing the adoption of generative AI and attempting to introduce it into communication tools. This paper introduces some aspects of those discussions and a prototype.

1. はじめに

コロナ禍に伴って遠隔地を繋いでビデオ会議や協働作業を行うツールが発展し、組織内や組織を跨いだミーティングなどにも活用されている。総務省の

情報通信白書令和 5 年版 [1] によると、株式会社東京商工リサーチが実施した調査において、最初の緊急事態宣言前のテレワークの実施率が 25.3% であったのに対し、2022 年以降は 30% 前後で推移しており、

一定の増加が見られる。対外的なミーティングやセミナーにも活用されており、帝国データバンクの調査 [2] では、50.2%の企業が主にハイブリッド、14.8%の企業が主にオンラインで社外との会議を実施していると答えている。会議型のツールだけでなく、ホワイトボードをインターネット上で共有してアイデア出しや整理などのグループワークを行えるオンラインホワイトボードツールも普及しており、様々な共同作業のオンライン化が一般的になっている。本稿では、これらビデオ会議やホワイトボードも含めた、オンラインで意思疎通を図り共同作業を行うためのツールをコミュニケーションツールと呼ぶ。

コミュニケーションツールの機能には様々なものがあり、議事録の作成など AI を活用した機能の導入も試みられてきた。ここ数年、生成 AI と呼ばれる、テキストや画像、音声、映像などのデータを、人間が作成したものに近い品質で生成する技術が発展してきている。特にテキストから画像を生成する Text-to-image と呼ばれる技術が進歩し、描画したい物の名前やスタイルをテキストで入力することで、品質の高い画像が生成できる。このような技術には、Dall・E[3] や StableDiffusion[4] が挙げられる。また、技術の進歩とともに、実装やモデルがオープンソース化され、より自由な条件で入手できるようになったことも、進歩をさらに後押ししている要因と考えられる。同様に、GPT-3[5] に代表されるテキスト生成技術や、従来から研究されている音声合成技術、またこれらの情報のメディアの種類を複数組み合わせることが可能なマルチモーダルな技術も発展している。マルチモーダル生成 AI の例

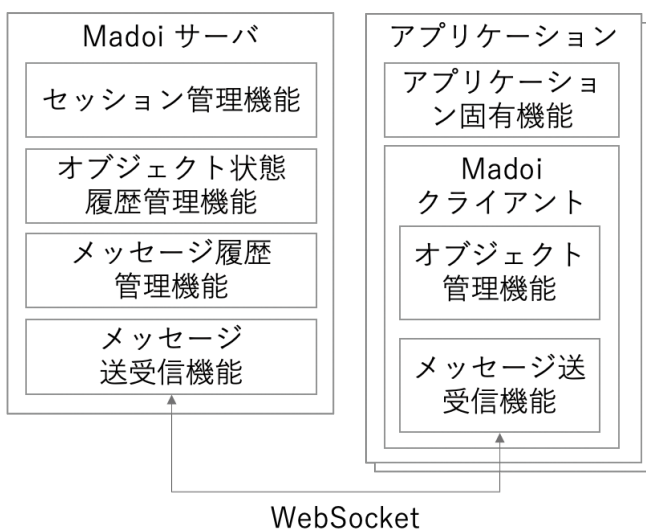


図1 Madoi を利用したシステムの構成

としては、Meta 社の ImageBind[6] が挙げられる。ImageBind では、画像から画像の説明文、音声から画像の生成など、複数のメディア形式を相互に変換する形で生成を行う。

未来環境ラボでは、コミュニケーションツールの未来像を探るべく、コミュニケーションツールのプロトタイピングや、プロトタイプ作成に必要な基盤機能の研究開発を行ってきた。コミュニケーションツールを作成するためには、まずネットワークを通じて情報や操作を共有するための仕組みが必要である。本稿では最初に、この情報共有を実現するフレームワーク、Madoi を説明する。次に生成 AI を取り込むためには、様々な生成 AI を容易に呼び出し結果を受け取るための仕組みが必要となる。そのため、MLServices と呼ぶ生成 AI のウェブサービスフレームワークを研究開発しており、これについて解説する。最後に、コミュニケーションツールのプロトタイプである Presence において、Madoi を用いた協働作業機能や MLServices を用いた生成 AI 機能を追加する試みについて紹介する。

2. 状態共有を実現する仕組み

オンラインコミュニケーションツールのプロトタイピングを行うにあたり、未来環境ラボではまず、アプリケーション内のオブジェクトの状態を共有する仕組み、Madoi の実現を試みた。Madoi は分散共有オブジェクトを実現する仕組み [7][8] をベースに実装されたものであり、アプリケーション内のオブジェクトの状態変更を生じさせる操作を、インターネットを跨いで共有することで、状態の共有を実現する。図 1 に Madoi を利用したシステムの構成を示す。Madoi の実装は、サーバとアプリケーション内のクライアント部分に分かれている。サー

```

window.addEventListener("load", ()=>{
  const m = new Madoi("room1");
  const chat = new Chat();
  m.register(chat, [{ // 方法1
    method: chat.addLog,
    share: {maxLog: 1000}
  }]);
});
class Chat{
  @Share({maxLog: 1000}) // 方法2
  addLog(name: string, message: string){
    ...
  }
}

```

図2 Madoi での共有メソッドの指定

サーバはセッションの管理とオブジェクト状態やアプリケーションとの間で送受信されたメッセージの履歴管理、およびメッセージ送受信機能から構成される。クライアントはサーバとのメッセージ送受信機能とオブジェクト管理機能から構成される。

サーバ側で管理するセッションの情報には、クライアントの接続状態、オブジェクトの状態に関するメッセージの履歴が含まれる。サーバ側では、これらの情報をもとに、アプリケーションに送信するメッセージを制御する。例えばクライアントがサーバに接続すると、そのクライアントにオブジェクトの状態やメッセージの履歴を送信し、他のクライアントには新しいクライアントが参加したことを通知する。オブジェクトに関する情報は、状態と変更操作の履歴から構成されており、ある時点での状態と、その状態以降の変更操作履歴を保持している。クライアントから状態が送信されてくると、状態を更新し、変更履歴を破棄する。これにより、変更操作履歴を全て保持する場合に比べて、少ない容量でオブジェクトの状態を記録できる。オブジェクトの状態管理はオプションな機能であり、これを利用しない場合は、汎用のメッセージ配送システムとして利用することもできる。

クライアント側では、アプリケーション内で共有の対象としているオブジェクトの管理や、サーバとのメッセージ送受信を行う。オブジェクトは登録順

でIDが採番され、あるIDを持つオブジェクトへの変更操作が、他のアプリケーション内で同じIDを持つオブジェクトにも適用される。変更操作の検出には、「変更操作を実装したメソッド」を「実際には変更操作を実行せずサーバに通知するメソッド」で置き換える仕組みを導入している。変更操作を実装しているメソッドを指定する方法は、2種類用意されている(図2)。方法1では、共有するメソッドをregisterメソッドの引数により指定している。メソッド名が変更された場合は指定内の記述も変更する必要があり、またメソッド定義を見ただけではそれが共有対象かどうか分からない。方法2では対象のメソッドに直接@Shareというアノテーションを指定している。この方法では、メソッド定義を参照すれば共有対象であることがわかる。

3. 生成 AI を利用する仕組み

次に、生成 AI をアプリケーションから利用する仕組み、MLServices[9][10]について説明する。図3に、MLServicesを利用する際のシステムの構成を示す。

コミュニケーションツールの研究を進めるにあたって、未来環境ラボでは、離れた場所にいる人同士を繋いで行う共同作業に注目してきた。近年の生成 AI の発展により、これまでコンピュータには難しいとされていた作業にも AI が活用されてきている。それは共同で行う作業も同様であろうと考え、その推測の下、多種多様に進化している途上の生成 AI を、コミュニケーションツール上で利用可能にする仕組みを検討するに至った。生成 AI を活用するにあたって直面する技術課題の一つは、実行に必要な計算能力の問題である。生成 AI の多くはニューラルネットを発展させた仕組みで実現されており、その実行に専用の GPU (Graphics Processing Unit) を必要とする。GPU は 3 次元仮想空間を描画するため、3 次元の座標系に存在する光源、カメラ、物体とその素材などの情報から、ディスプレイの各点に描画すべき色を求めるために行列計算を行う。行列計算の速度向上が描画性能の向上に直結するため、GPU は行列計算が高速に行えるように発展してきたが、ニューラルネットの実装方法として行列計算を行う実装が一般的になったことで、GPU を用いた学習・推論手法が発展した。生成 AI でも GPU を利用するものが多く、推論の速度は GPU の性能に依存する。しかし、高性能な GPU



図3 MLServices を利用したシステムの構成

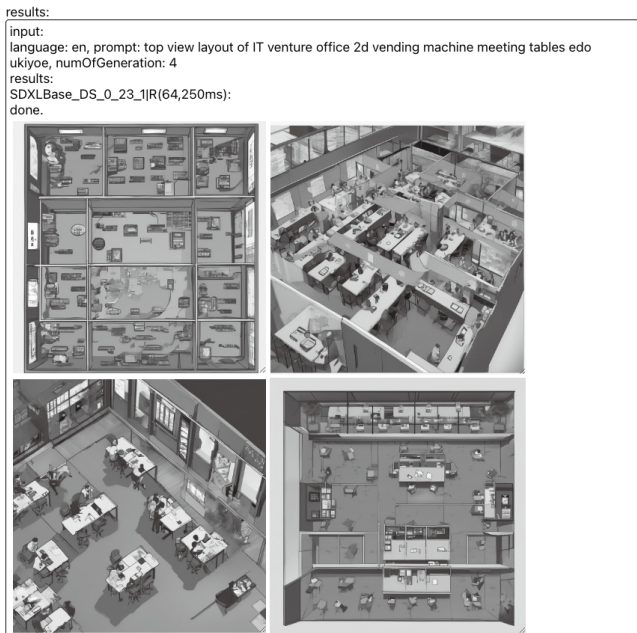


図4 MLServicesの動作確認用UI

は消費電力や発熱量が多いため、利用者のPCに搭載されるGPUの性能は限られる。そのため、より多様な生成AIを利用可能にするには高性能なGPUを搭載したサーバ側で処理を実行し、それを利用者のPCから利用できる仕組みが必要である。また、生成AIはさまざまな実装が存在しており、実装毎に利用方法が異なるため、利用時に個別にプログラムを作成する必要がある。プロトタイピングを行うには多種多様な生成AIを切り替えて試す必要があり、個別に対応するとコストがかかる。

MLServicesでは、高性能なGPUの利用と個別の生成AIへの対応コスト削減を目的として、サーバ側で動作する機能をクライアントから利用する、Webサービス化を行なっている。MLServicesはサーバ側の機能とクライアント側の機能から構成され、サーバ側では、生成AIの種類に応じた呼び出しインターフェースの定義、生成AIへのリクエストとレスポンスを効率よく送受信するためのプロトコルの定義と実装、生成AIのプロセスやサービス定義を管理するための仕組みの定義と実装、動作確認用UIの実装を行なっている。クライアント側では、サーバを経由して生成AIを実行し結果を受け取るためのクライアントライブラリを提供している。

インターフェース定義としては、画像生成や対話、音声認識、音声合成、形態素解析など、39のサービスインターフェースが用意されている。プロトコルでは、通信の即時性とデータ容量の削減のため、WebSocketを用いたバイナリプロトコル上に、サー



図5 Presenceの画面

ビスインターフェースに応じたリクエストおよびレスポンスが送受信できるプロトコルが実装されている。また、これらインターフェースとプロトコルに対応した、各サービスの動作確認用UIが用意されている(図4)。MLServicesに登録されたサービスは、この動作確認用UIにウェブブラウザからアクセスし、実行結果を確認できる。

生成AIをウェブサービス化するための仕組みは、サービス管理機能とプロセス管理機能から構成される。サービス管理機能は、あらかじめ用意された定義ファイルを読み込んでサービスを登録する。サービスの情報は、サービス名、説明、URL、ライセンス状態、実装情報から構成される。実装情報以外の情報は、管理APIを通じてクライアントから取得できる。実装情報は、サーバ内でそのサービスの実行を担当するクラス名、実行されるプロセスの起動コマンド、GPUの必要メモリサイズから構成される。プロセスは起動コマンドで識別される。サービス実行があった際に、同じ起動コマンドで起動されたプロセスがあれば再利用され、無ければ新たに起動される。サービス実行のリクエストおよびレスポンスは、標準入出力を通じて送受信される。起動コマンドが必要とするGPUメモリサイズと、サーバマシンに用意されているGPUメモリサイズを元に、同時にプール内に存在できるインスタンスが制限される。サーバマシンのGPUメモリが許す範囲でインスタンスがプールされるため、同じプロセスの実行が連続する場合、プロセスの起動時間が短縮され、実行時間が削減できる。この仕組みを用いて、現在、280のサービスが用意されている。また、MLServicesのソースコードはGitHubで公開されている¹⁾。

1) <https://github.com/openlangrid/mlgrid-services/>

4. 生成 AI を活用したコラボレーションツール

未来環境ラボではこれらの機能を活用し、共同作業のためのツールを備えたビデオ会議システムである Presence のプロトタイピングを行なっている。図 5 にその画面例を示す。Presence では、ビデオ会議に必要な音声と映像の送受信機能に加えて、左下に仮想空間内での互いの位置を共有する機能、右下にさまざまな共同作業ツールを備えている。チャットやホワイトボードのほか、画像生成 AI による生成結果を共有するツールや、大規模言語モデル (LLM: Large Language Model) によるテキスト生成の結果を共有するツールも備えている。チャットでは、形態素解析により抽出した単語の頻度からワードクラウドを描画する機能や、チャットの内容を定期的に LLM で要約する機能も実装されている。これらツールの共有部分は、Madoi を用いて実現されている。生成 AI を用いて実現している要約、画像生成、テキスト生成は、MLServices を通じて利用できるものを使用している。

コミュニケーションツールが備える機能は、遠隔地の利用者同士のコミュニケーションを実現するための機能と、作業内容を共有する機能に分けられる。前者は音声と映像をやりとりすることで、利用者の表情や動作を互いの視覚と、声を互いの聴覚と接続し、コミュニケーションを促す。後者はコンピュータを使用して行う文書作成作業や描画作業、スケジュールの作成や表計算などがあり、それぞれの用途に用意されたアプリケーションやツール内の特定の状態を共有することで、共同作業を可能にする。AI はこれらの機能に対して組み合わせられ、前者では声を認識して文字起こしを行い、内容の翻訳や要約を行う。また、カメラ映像の加工にも利用される。後者では、文書作成や絵の描画などの作業そのものの一部を生成 AI が担うようになってきている。コミュニケーションツールでは、このような変化に対応し、さらに多様な機能を共同で実施できるようにすることや、作業そのものを AI に代替させることが求められている。

5. おわりに

本稿では、コミュニケーションツールのプロトタイプを作成するために必要なツール、アプリケーションを跨いだ情報共有を実現する Madoi、およ

び生成 AI をウェブサービスとして提供・利用する MLServices について説明し、実際にそれらを利用してプロトタイピングした Presence を解説した。インターネットの発達とともに様々なコミュニケーションツール・コラボレーションツールが研究され実用化されてきたが、新型コロナウイルスの流行を契機にネットワークを跨いだ共同作業の需要が急速に増加し、今後も一定の需要が見込まれる。ビデオ会議システムについては一般的になったが、人間がコンピュータを利用して行なっている作業は様々なものがあり、また未だ IT 化されていない作業も含めて、コミュニケーションツールが実装すべき共同作業機能の候補は今後も増加すると見込まれる。さらに、生成 AI の発展、普及によって作業そのもののやり方も変わろうとしている。このような状況を踏まえ、未来環境ラボでは、コミュニケーションツールの未来像を見据え、また、生成 AI による変革も考慮しつつ、コミュニケーションツールの研究として今後も考察とプロトタイピングを行なっていく。

参考文献

- [1] 総務省, “令和 5 年版情報通信白書”, 2023.7, URL: <https://www.soumu.go.jp/johotsusintokei/whitepaper/index.html>, 2024.1.25 取得
- [2] 帝国データバンク, “社内外会議に関する企業の実態調査”, 2023.4, URL: <https://www.tdb.co.jp/report/watching/press/pdf/p230409.pdf>, 2024.1.25 取得
- [3] Ramesh, Aditya, et al. "Zero-shot text-to-image generation." International Conference on Machine Learning. PMLR, 2021.
- [4] Rombach, Robin, et al. "High-resolution image synthesis with latent diffusion models." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022.
- [5] Brown, Tom, et al. "Language models are few-shot learners." Advances in neural information processing systems 33 (2020): 1877-1901.
- [6] Girdhar, Rohit, et al. "Imagebind: One embedding space to bind them all." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023.
- [7] 中口 孝雄, “リアルタイムコラボレーションツールのためのオブジェクト共有サービス”, 電子情報通信学会技術研究報告; 信学技報 120.232, pp.70-73, 2020.11.14.
- [8] 中口 孝雄, “分散共有機能のサービス化に向けたプログラミングモデルの設計と実装”, 電子情報通信学会技術研究報告; 信学技報 121.157, pp.64-68, 2021.8.20.
- [9] 中口 孝雄, “機械学習システムへの複合サービス技術の適用”, 通信学会技術研究報告; 信学技報 119.178, pp.39-40, 2019.8.16.

[10] 中口 孝雄, “機械学習サービスを登録・提供するサービス基盤の構築に向けて”, NAIS Journal, vol.15, pp.66-73, 2021.6.

◆著者紹介

中口 孝雄 Takao Nakaguchi

未来環境ラボ
京都情報大学院大学 准教授
株式会社コネクションズ 代表取締役
京都大学 大学院情報学研究科 博士 (情報学)
京都情報大学院大学 情報技術修士 (専門職)
京都コンピュータ学院鴨川校 専門士

秋山 功 Isao Akiyama

未来環境ラボ
BIPROGY 株式会社 総合技術研究所 担当部長
早稲田大学 工学士

三浦 仁 Hitoshi Miura

未来環境ラボ
BIPROGY 株式会社 総合技術研究所 主席研究員
富山大学 工学士

前納 一希 Kazuki Maeno

未来環境ラボ
京都コンピュータ学院
京都コンピュータ学院京都駅前校 専門士
龍谷大学 文学士

橋本 昇 Noboru Hashimoto

未来環境ラボ
京都情報大学院大学 SOIC セクレタリー・ゼネラル
前, 日本ユニシス株式会社 (現, BIPROGY)
プロジェクト・マネジメント・スペシャリスト

江尻 秀彰 Hideaki Ejiri

未来環境ラボ
BIPROGY 株式会社 総合技術研究所

コラム

蘆山寺へ

編集子

紫式部 (970?-1019?) ゆかりの地である蘆山寺は寺町通りの今出川と広小路の間にあります。元々は彼女のひいお爺さんである藤原兼輔の地だったのを受け継いだそうです。藤原道長 (966-1028) の土御門邸も近くにあって、そこへは安倍晴明 (921?-1005) や源頼光 (948-1021) も出入りしていたでしょう。また時代は下りますが藤原定家の京極一条邸跡もこの近くで石碑が建っています。

なお当時の内裏の範囲は今の千本通りを中心に東は大宮通りまででした。

蘆山寺の門を入ると歌碑があり紫式部と娘の賢子 (大式三位: だいにのさんみ 999-1082) の歌が刻んであります。どちらも百人一首に載っている有名な歌です。

母紫式部の歌は

めぐりあひて見しやそれともわかぬまに 雲がくれにし夜半の月かな
「幼馴染と久しぶりに会ってすっかり夜半まで話しこんでしまった」
というわかりやすい素直な歌です。

ところが娘の歌は

ありま山みなの篠原風吹かばいてそよ人をわすれやはする

読者諸賢を前に解釈はできませんが、言い寄る元カレをそれとなく袖に振るというコワイ歌なのです。

母が仕えた彰子中宮の孫となる後冷泉天皇の乳母となり、即位とともに従三位に昇進し 80 歳過ぎまで生き延びます。歌の道も恋の道も世渡りの道も母親以上だったように思えますね、ひよっとしたら源氏物語は母娘の合作か?

清少納言, 紫式部母娘, 和泉式部母娘, 赤染衛門, 伊勢大輔などなど当時の平安京には世界に類を見ないほどあまたの才女が活躍していました。 「あすとろん」 62 号 p18 より抜粋