

人工知能（AI）と共生する言語教育を考えるシンポジウム

大規模言語モデルを用いた 外国語活動支援ツールの設計と開発

Version Feb-03-2024

長谷部 陽一郎

同志社大学

自己紹介

- 長谷部 陽一郎（はせべ よういちろう）
 - 所属：同志社大学グローバル・コミュニケーション学部
 - 専門分野：認知言語学、コーパス言語学
 - ウェブサイト：<https://yohasebe.com>
 - GitHub：<https://github.com/yohasebe>
 - メール：yhasebe@mail.doshisha.ac.jp

自己紹介

- これまで作ってきたもの
 - [TED Corpus Search Engine \(TCSE\)](#)
言語研究・言語教育用TED Talkコーパス検索システム
 - [RSyntaxTree](#)
言語研究用樹形図作成システム
 - [jReadability / jWriter](#) (李在鎬さんとの共同プロジェクト)
日本語文章難易度判別 / 日本語学習者作文評価 システム
- 本日紹介するシステム
 - [Monadic Chat](#)
外国語学習や教育での利用を想定したAIチャットボット・システム

※ MacとWindowsでそれなりに動く状態ですが、開発途上のシステムです。
思わぬ不具合が生じることもあります。細かい使用は結構頻繁に変わります 

本発表の目的

- 外国語科目を実際に担当する教員としての問い
 - 大規模言語モデルを外国語の授業にどのように活用できるか
 - どうすればそれを実現できるか
- 外国語の授業で使えるシステム（Monadic Chat）の開発
 - ChatGPTと同じOpenAIの大規模言語モデルを利用している
 - MacとWindowsで動作するシステムをオープンソースで公開
- AIと共生する言語教育
 - 現時点ではAIは個人が「こっそり使うもの」（？）と認識されているようにみえる
 - 教師が教室内での活動の幅を広げられるツールとしての可能性を模索したい

※ 同志社で全学の英語を教えています、
本発表で示すシステムを担当しているクラ
スを超えて広範囲に展開しているわけでは
ありません

無料で使えます。ただし OpenAI
の言語モデルの使用料は必要です

発表の内容

1. 大規模言語モデルを用いた外国語活動支援について
2. 教室内の外国語活動に対する支援の可能性
3. 教室内の外国語活動支援システムに求められるもの
4. どのようなツールを組み合わせるか
5. Monadic Chat のインストール／何ができるか／設定・表示項目
6. Monadic Chatを用いた教室内活動の提案
7. まとめと課題
8. 質疑応答

大規模言語モデルを用いた外国語活動支援について

大規模言語モデル（LLM）の活用の可能性

教師支援の例

- 長文読解での文章レベル判定
- 学習者による作文の添削
- 単語リストやクイズの作成
- 素材テキストの難易度調整

学習者支援の例

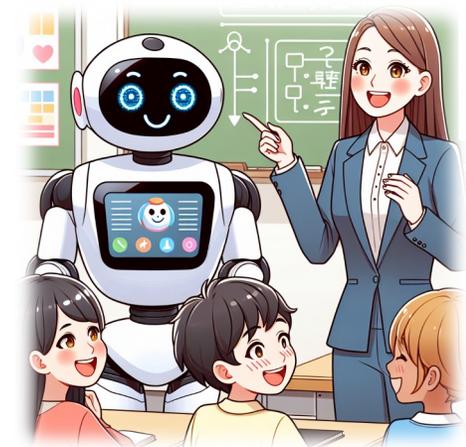
- 長文読解での訳と解説の提示
- 作文での言い換え表現の確認
- スピーキング練習
- 読み上げ音声の出力

上記はどちらかということと教室外もしくは個々人での活動

教室内でのリアルタイムの外国語活動に対する支援ができないか？

教室内の外国語活動に対する支援の可能性

- 教室内に音声によるAIとの会話を導入する
 - リアルタイムでの「会話」は、外国語の教育・学習において重要なはず
 - AIとの会話で「チームティーチング」的な状況を作り出す
- 教室内での外国語によるインタラクティブなやりとりの支援
 - 背景的知識・補足事項の解説
 - 語句やイディオムの解説や例文の提示
- 知識や能力の拡張
 - 使用しているテキストやこれまでの授業で扱った内容から必要事項をすぐに取り出す
 - 教師にとって「できなくはない」が「うまくやるのが難しい」様々なことへの支援



教室内の外国語活動支援システムに求められるもの

- テキスト入出力

シンプルなテキスト入力と見やすいテキスト表示

- 教室内のプロジェクター／スクリーンの利用を想定

- 音声入力

- PC／Macのマイクを使って教師や学習者の音声発話を認識できること

教室内が完全に無音になることはないのでノイズ耐性が求められる

- 音声出力

- 教室内のスピーカーから（なるべく自然な）音声出力ができるように

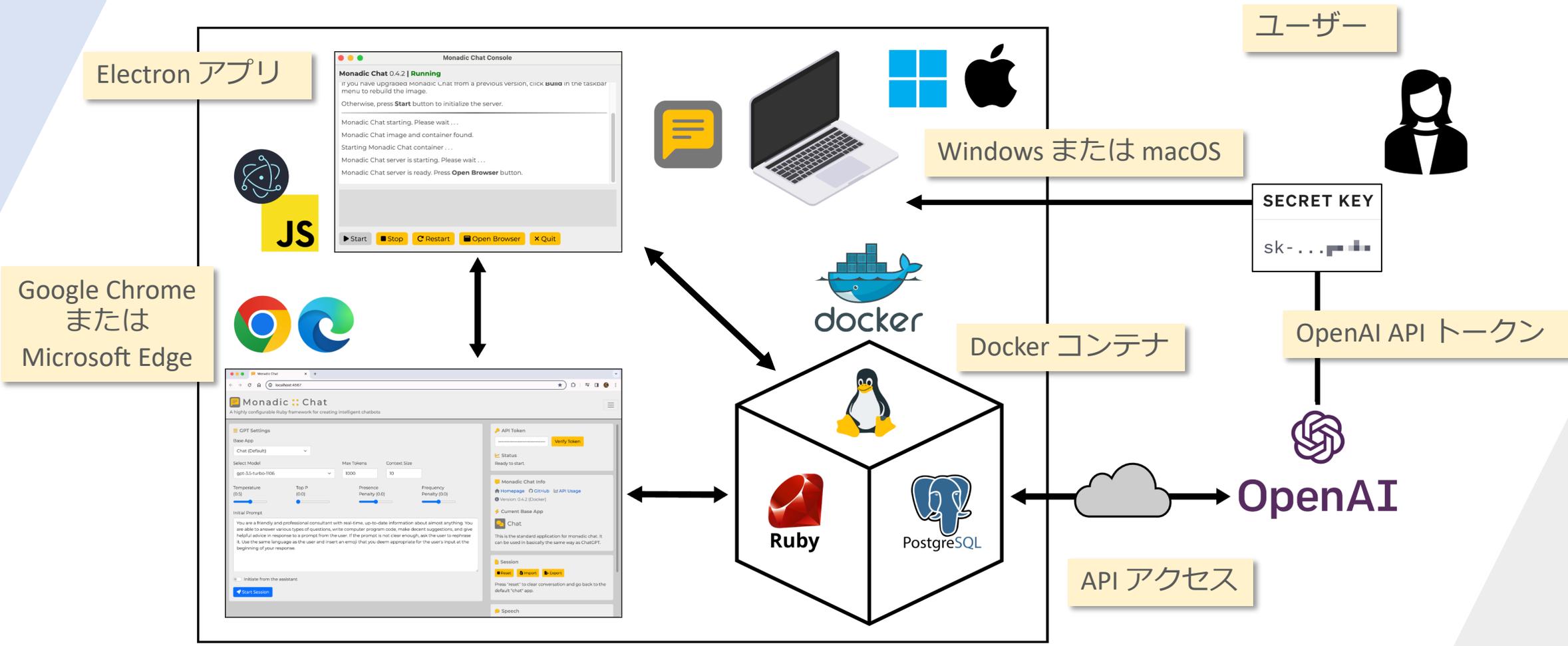
問いかけから返答までのタイムラグはなるべく減らしたい

- 画像ファイル入力

- 写真や図版などを読みこめることが望ましい

活字や手書き文字も入力・認識できるとなお良い

Monadic Chat の設計



どのようなツールを組み合わせるか (UI)

ユーザーインターフェイス (UI)

- ウェブブラウザを使用 Google Chrome または Microsoft Edge
- 誰もが使い慣れている
- 多言語のテキスト入出力が可能
- 音声や画像の入出力が可能 教室内でコンピュータに張り付く状況はできれば避けたい
- 各種入力装置（キーボード、マウス、その他コントローラー）が使用可能
- 複数のOSへの対応が容易 Windows (x64)、macOS (x64, Intel)、macOS (Apple Silicon)

どのようなツールを組み合わせるか（API）

大規模言語モデルAPIの利用

- OpenAIの各種モデルのAPIを利用
 - チャット／テキスト補完：gpt-3.5-turbo、gpt-4、gpt-4-turbo-preview
 - 画像認識：gpt-4-vision-preview
 - 画像生成：dall-e-3
 - 音声合成：tts-1
 - 音声認識：whisper-1
 - テキスト埋め込み：text-embedding-3-small

API = application programming interface

ユーザーが取得した[OpenAIのAPIキー](#)を使用
([有料](#)、ChatGPT Plusの契約とは別。)

GPT 3.5系列のモデルはアカウント作成後すぐに利用可能。[GPT 4系列のモデル](#)は、アカウント作成後に\$1以上の請求・支払が完了した後に利用可能になる。

どのようなツールを組み合わせるか（統合）

Docker による仮想環境の利用

- 利点

- Linuxの仮想環境を構築してPC内で安全にサーバーを起動できる
- ほぼ同一のコードでシステムをWindowsとMacに対応させることができる
- 内部で Vector DB を稼働させて RAG を行うことができる（後述）
- ローカル環境を汚すことなくインストール／アンインストールできる

- 欠点

- インストールに手間と時間がかかる 手順通りやればそんなに難しくない（はず）
- WindowsではWSL2（Windows Subsystem for Linux）もインストールが必要
- Dockerの「イメージ」や「コンテナ」といった概念がわかりづらい

わからなくても利用可能

Monadic Chat のインストール

• macOS

1. Docker Desktop
 - [インストーラ](#)をダウンロード
 - すべて標準の設定でインストール
2. Monadic Chat
 - [インストーラ](#)をダウンロード
 - Intel 版
 - Apple Silicon 版
 - インストーラを実行して、アプリを Application フォルダに移動

Intel 版 と Apple Silicon 版 の区別にご注意ください

• Windows

1. WSL2
 - Windows 11 の場合、Microsoft Store からインストール可能
2. Docker Desktop
 - [インストーラ](#)をダウンロード
 - すべて標準の設定でインストール
3. Monadic Chat
 - [インストーラ](#)をダウンロード (64bit版)
 - インストーラを実行

Windows 版のインストール時には警告が出るかも・・・
「詳細情報」をクリックして「実行」してください。
インストール実績が一定以上になるまでの制約らしい。

Monadic Chat の起動と利用

初回起動時とアップデート直後はイメージのビルドが必要

1) イメージのビルド → 2) コンテナ起動 → 3) ブラウザでアクセス

MacBook Pro (2021 M1) の場合、1分半程度でイメージのビルドとコンテナの起動が完了

• 初回起動時

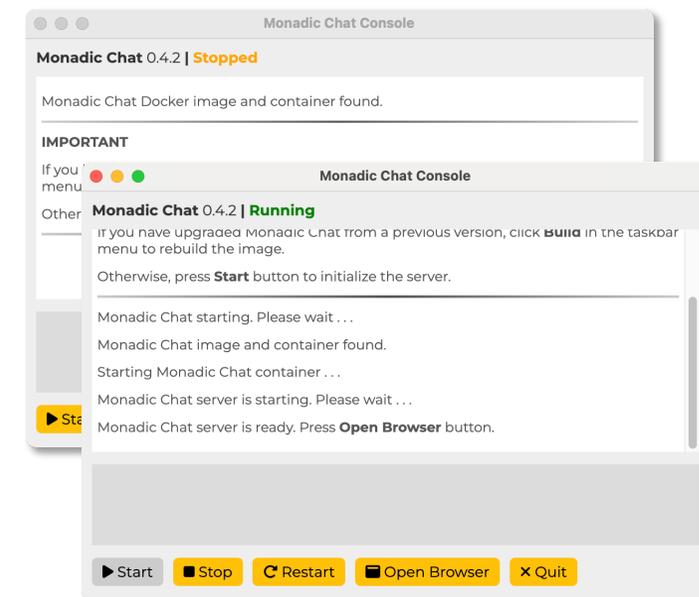
- Start ボタンでビルド開始 → 自動的にコンテナ起動
- Open Browserでブラウザを立ち上げ

• バージョンアップデート直後

- “メニュー / Build” で再ビルド → Start ボタンでコンテナ起動
- Open Browserでブラウザを立ち上げ

• その他の場合

- Start ボタンでコンテナ起動 → Open Browserでブラウザを立ち上げ



Monadic Chat の設定・表示項目

GPT Settings

ベースアプリ選択

Base App
Chat (Default) ▾

モデル選択 + 各種パラメーター設定

Select Model
gpt-3.5-turbo-1106 ▾

Max Tokens: 1000
Context Size: 10

Temperature (0.5) [Slider]
Top P (0.0) [Slider]
Presence Penalty (0.0) [Slider]
Frequency Penalty (0.0) [Slider]

Initial Prompt
You are a friendly and professional consultant with real-time, up-to-date information about almost anything. You are able to answer various types of questions, write computer program code, make decent suggestions, and give helpful advice in response to a prompt from the user. If the prompt is not clear enough, ask the user to rephrase it. Use the same language as the user and insert an emoji that you deem appropriate for the user's input at the beginning of your response.

Initiate from the assistant

基本プロンプト編集

Start Session

API トークン入力

API Token
[Input Field] **Verify Token**

Status
API token verified and stored in .env file.

利用可能モデル

Available Models
gpt-4-turbo-preview
gpt-4-0125-preview
gpt-3.5-turbo-1106
gpt-4-1106-preview
gpt-4-vision-preview
gpt-4
gpt-3.5-turbo-16k
gpt-3.5-turbo

セッション管理

Session
Reset **Import** **Export**

Press "reset" to clear conversation and go back to the default "chat" app.

音声合成設定

Speech
Voice: Alloy ▾
Speed (1.0) [Slider]

Monadic Chat Info

[Homepage](#) [GitHub](#) [API Usage](#)

Version: 0.4.2 (Docker)

Current Base App

Chat

This is the standard application for monadic chat. It can be used in basically the same way as ChatGPT.

ベースアプリ情報

Dialog

Use easy submit (enter key or stop button) Auto speech

Role: User ▾

Type your message...

Press **Send** to send your message.

メッセージ入力

Send **Clear** **Use Image** **Voice Input**

画像アップロードと音声入力

Monadic Chat でできること

利用可能なベースアプリ（主なもの）

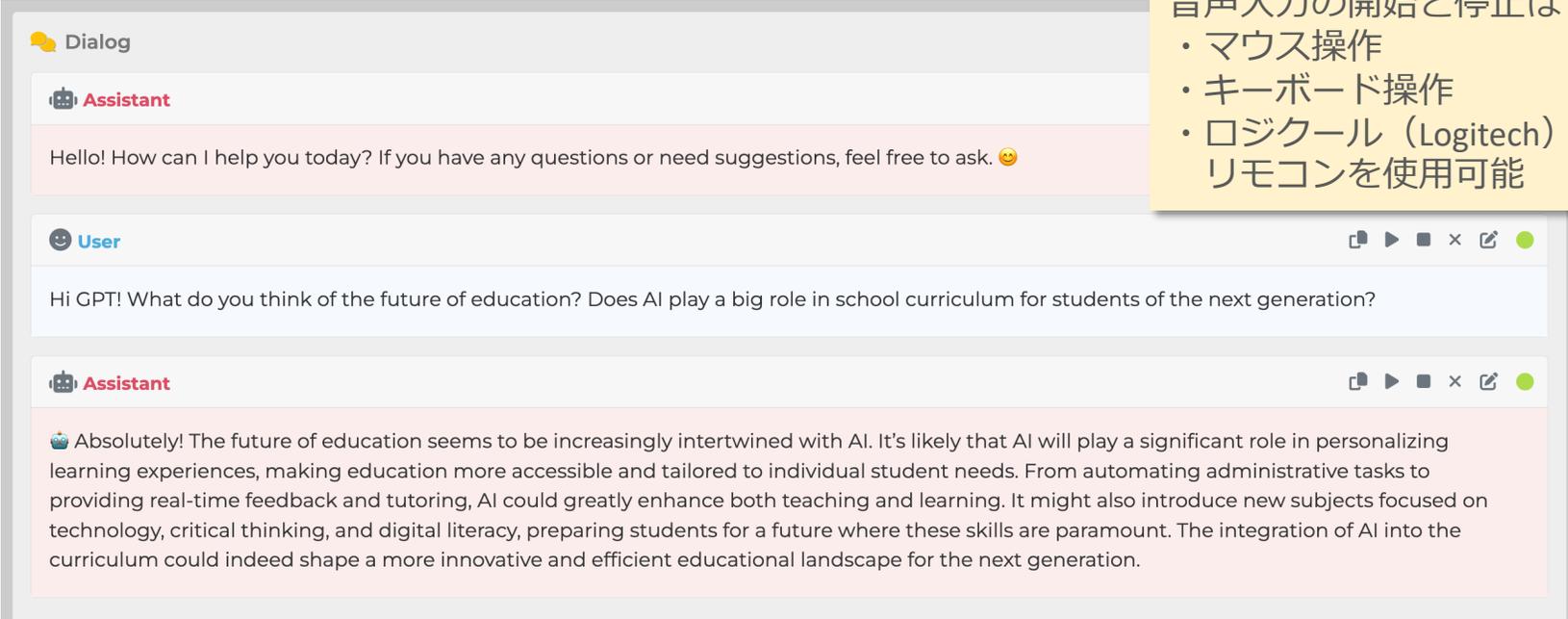
言語教育・学習での利用を常に念頭に置いていますが、それだけにとどまらない汎用的なシステムとなっています。

- Chat テキストベースのチャットを行う標準アプリ
- Code コンピュータ・プログラム開発支援
- Image Generator チャット形式で画像生成
- Language Practice Plus 音声認識・音声合成による外国語学習
- Mail Composer Eメールのメッセージ代筆
- PDF Navigator PDFをアップロードしてRAGを用いたチャット
- Translate テキスト翻訳
- Voice Chat / Voice Interpreter 音声認識・音声合成によるチャット / 通訳

教室内活動の提案：音声機能を用いて

例1：教室内で話題に上ったトピックについて、AIに説明や意見を求めてみる

例2：事前にプロンプトで設定を与えておき、授業の内容に即したロールプレイをさせる



Dialog

Assistant

Hello! How can I help you today? If you have any questions or need suggestions, feel free to ask. 😊

User

Hi GPT! What do you think of the future of education? Does AI play a big role in school curriculum for students of the next generation?

Assistant

Absolutely! The future of education seems to be increasingly intertwined with AI. It's likely that AI will play a significant role in personalizing learning experiences, making education more accessible and tailored to individual student needs. From automating administrative tasks to providing real-time feedback and tutoring, AI could greatly enhance both teaching and learning. It might also introduce new subjects focused on technology, critical thinking, and digital literacy, preparing students for a future where these skills are paramount. The integration of AI into the curriculum could indeed shape a more innovative and efficient educational landscape for the next generation.

Send Clear Use Image Stop

音声入力の開始と停止は

- ・マウス操作
- ・キーボード操作
- ・ロジクール (Logitech) のプレゼンテーション・リモコンを使用可能



OpenAIの音声認識と音声合成の特徴

音声認識

- 音声認識モデルWhisperは周囲の雑音に強い。多くの場合、教室内でも問題なく使える。
- 英語や日本語を含む多言語に対応。
- Googleの音声認識エンジンなどとは異なり、認識の「確らしさ」の値を得ることはできない。

→ 発音の正確さを測定する用途には使えない。

- 文法的に問題のある発話であっても（動詞の活用や名詞の単複など）適宜修正して表示することがある

音声合成

- 現時点では英語に最適化されているが、多言語に対応している。
- 日本語などの言語でも利用可能だが、やや不自然さがある。

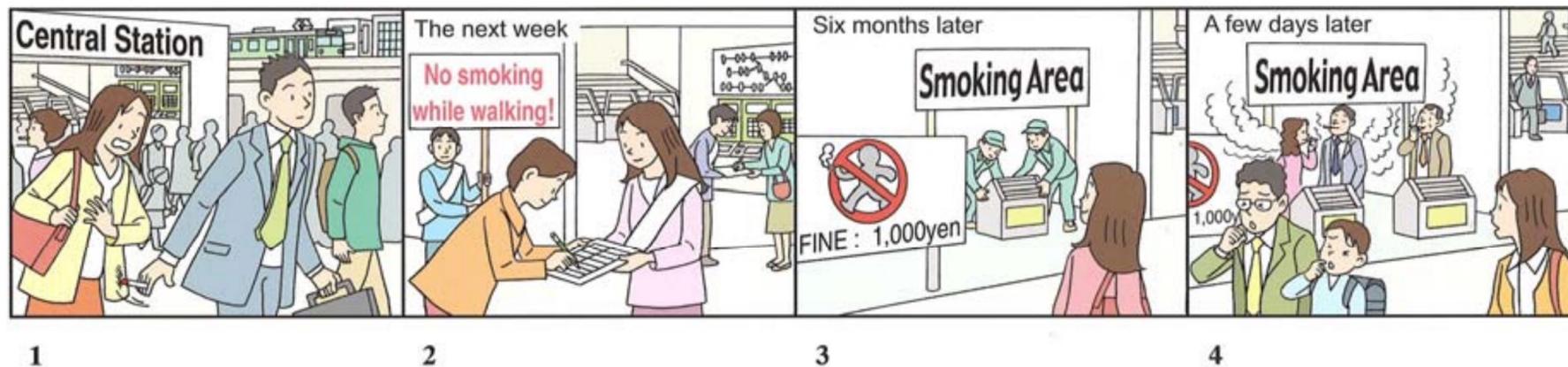
→ ちよくちよく漢字を読み間違える。

- 6種の音声を利用可能
- モデルとなっている人物や英語のバリエーションに関する情報などは明らかにされていない

教室内活動の提案：画像認識を用いて

例1：身近な風景や状況を示した画像を用いて対象言語での説明を試みる

例2：TOEICの写真描写問題や英検2次面接のイラスト説明対策



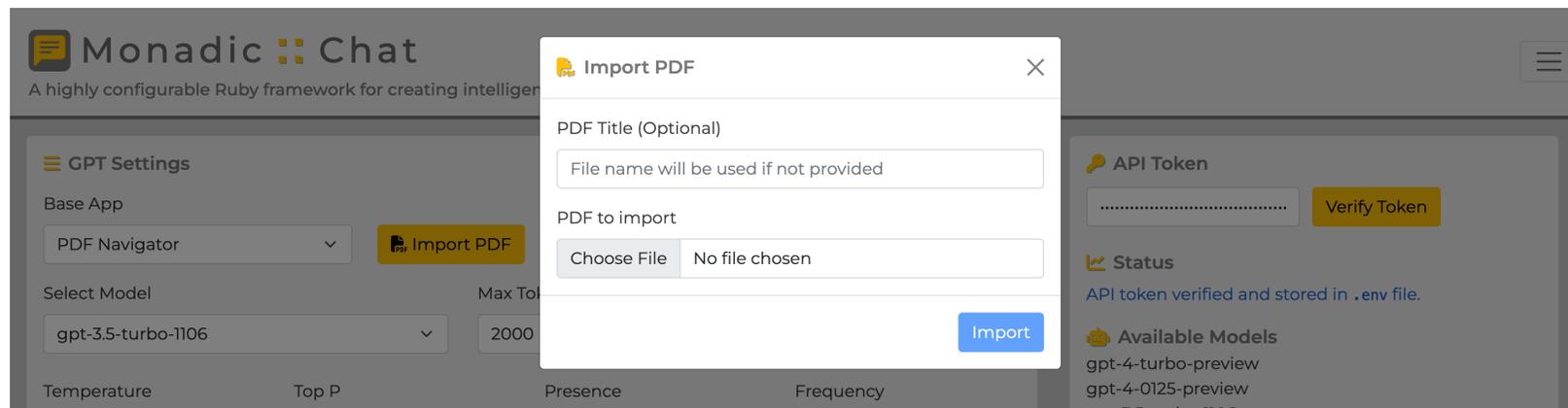
英検準1級2次試験サンプル問題より

https://www.eiken.or.jp/eiken/exam/virtual/grade_p1/pdf/grade_p1.pdf

教室内活動の提案：検索拡張生成（RAG）を用いて

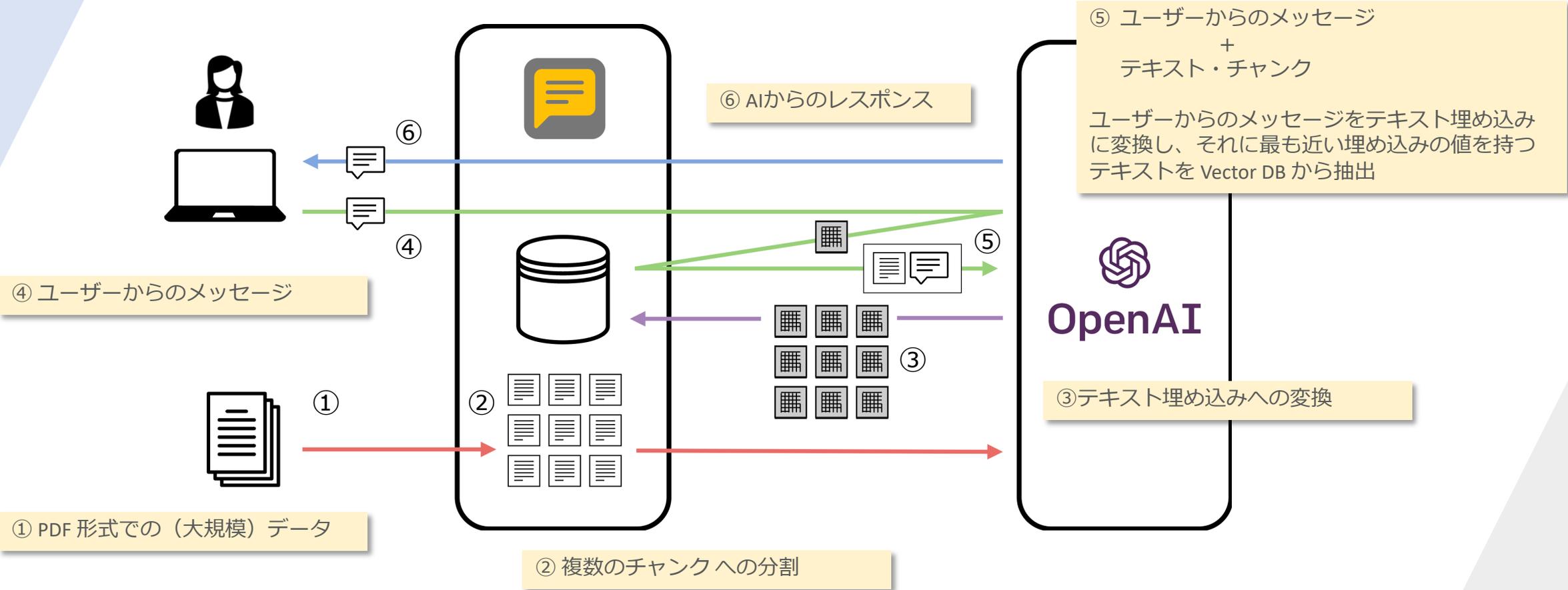
例1：授業で使う資料の全体を事前に読み込ませて、関連する質問をする

例2：授業で使う用語の定義や索引情報などを読み込ませて、適宜教室内で参照する



ユーザーからのメッセージに関連する情報を適宜 Vector DB (pgvector) から取り出し、元のメッセージに添えてAPIへの問い合わせを行う。

検索拡張生成 (RAG=retrieval augmented generation)



今後の展望

- ユーザーからの発言に対するモデレーションの機能

OpenAIの [moderation API](#) は無料で使える

- OpenAI以外のモデルへの対応

中期的な展望

- ローカルで完結した仕組みの構築

長期的な展望

- ユーザーの発音の正確さ・自然さの評価

ブラウザの web speech API を使えばできるが . . .

- より高度な「コンテキスト」保持機能

Monadicという名前の由来なので . . .

コンテキストの保持について

OpenAIの大規模言語モデルとの対話が（現時点で）人間同士の会話と異なる点：

- 1回1回のリクエストで「コンテキスト」を与えてやる必要がある

コンテキストを共有しているように見えて、実はターンごとに「お膳立て」をしている

```
messages=[  
  {"role": "system", "content": "You are a helpful assistant."},  
  {"role": "user", "content": "Who won the world series in 2020?"},  
  {"role": "assistant", "content": "The Los Angeles Dodgers won the World Series in 2020."},  
  {"role": "user", "content": "Where was it played?"}  
]
```

お膳立てのやり方はいろいろ：

- 直近の発話ターン
- 選好発話の要約
- テキスト埋め込みデータ（RAG）

API レスポンス

```
response = {  
  "role": "assistant",  
  "content": "The 2020 World Series was played in Texas at Globe Life Field in Arlington."  
}
```

まとめと課題

• データの観点からみた課題

- 個人情報保護に関する問題
- 著作権保護に関する問題
- 大規模言語モデルが出力する内容の正確さ・適切さに関する問題

ChatGPTを使う場合と違い、APIを通じて入力されたテキストはOpenAIの大規模言語モデル開発のためのデータとして利用されないことが[明記](#)されています。

• システム開発の観点からみた課題

- 急速に発展している様々な機能サービスへの追従
- 現時点では頻繁に起こる OpenAI 側での API エラーの問題

OpenAIの各種サービスの稼働状況は[ここで](#)確認できます。

次々と新たなツールが開発される中で、また既存のツールが進化していく中で、手を動かしてシステム開発を続けることの意義とは？

参考資料

- Hasebe, Yoichiro. 2023. [Monadic Chat : テキスト補完 API で文脈を保持するためのフレームワーク](#). 『言語処理学会 第29回年次大会 発表論文集』
- Hornstein, Norbert. 2023. 「[人工知能という分野が謙虚であったことなど一度もない](#)」 折田奈甫・藤井友比呂・小野 創〈編訳〉『科学』2023年12月号. 岩波書店.
- Martineau, Kim. 2023. [What is retrieval-augmented generation?](#) IBM Research.
- OpenAI. 2023. [GPT-4 Technical Report](#). arXiv:2303.08774 [cs.CL]
- Ovadia, Oded, Menachem Brief, Moshik Mishaeli, and Oren Elisha. 2023. [Fine-Tuning or Retrieval? Comparing Knowledge Injection in LLMs](#). arXiv:2312.05934 [cs.AI]
- Radford, Alec, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. 2022. [Robust Speech Recognition via Large-Scale Weak Supervision](#). arXiv:2212.04356 [eess.AS]
- Torrent, Tiago, Thomas Hoffmann, Arthur Lorenzi, and Mark Turner. 2023. [Copilots for Linguists: AI, Constructions, and Frames](#). Cambridge University Press.
- Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. [Attention Is All You Need](#). arXiv:1706.03762 [cs.CL]