

AI 議事録導入事業に係る仕様書

1 事業目的

現在、本市では職員が議事録を作成する際、音声データの聞き戻し作業が発生する等、議事録作成作業に多くの時間が必要な状況です。

そこで、AI 音声認識技術を活用した議事録作成支援システムを導入することで、議事録等の作成にかかる時間を削減し、職員の事務作業の負担軽減と業務効率化を図るものです。

2 業務の概要

- (1) 会議録作成のため議事録作成支援システムをオンプレ版にて提供すること。
- (2) 議事録作成支援システムライセンスは、4ライセンスを提供すること。
- (3) 議事録作成支援システムの導入及び運用が円滑に行えるよう、初期設定等を速やかに行うこと。
- (4) 市に対して操作研修会^{*}を実施すること。

※市が指定した日で1日（午前・午後で1回ずつ）を想定

3 業務内容

別紙『システム機能要件書』のとおりとする。

なお、仕様書に記載のない事項であっても、業務を実施するために必要な事項は実施するとともに、これらの費用を負担しなければならない。

4 契約期間

契約締結日から令和7年6月30日まで

5 納入場所及び納入期限

<納入場所> 守口市役所デジタル戦略課

<納入期限> 令和4年7月25日

6 納品物

音声認識システム『AmiVoice ScribeAssist』4ライセンス

操作マニュアル 一式

運用マニュアル 一式

7 その他

- (1) 導入機器の円滑な運用のために市を支援するとともに、受託者からの調査依頼、資料請求等に対して迅速に対応すること。
- (2) 本業務の履行に伴い発生する成果物等はすべて市に帰属するものとする。
- (3) 業務の実施に当たりデータの漏えい、データの滅失、事故等の予防に十分留意し、業務の信頼性、安全性の確保に努めること。
- (4) 業務の総括責任者及び代行するものを置くこと。総括責任者は、業務実施中に従事者を指揮し、市の担当者と連絡を密にし、遺漏のないように努めること。

8 特記事項

- (1) 本システムの安定的な運用を行うため、運用マニュアルを作成し納品すること。
- (2) 業務の実施に当たって知りえた業務上の秘密を外部に漏らし、又は他の目的に利用してはならない。
- (3) AmiVoice ScribeAssist ライセンス利用に伴う月額費用は毎月払いとし、初期費用は初回の月額費用とあわせて支払うものとする。
- (4) 本仕様書の内容について疑義が生じた場合、又は本仕様書に定めのない事項については、市と協議の上決定する。

システム機能要件書

1. システム仕様

1-1. 基本機能

- (1) Web 会議や商談、会見等、対面・非対面の幅広いシーンで利用できるスタンドアローン型文字起こし支援アプリケーションを提供すること。
- (2) リアルタイムでの音声認識に加え、録音した音声データを取り込んで文字化するバッチ認識にも対応していること。
- (3) 音声認識で文字化されたテキストから順次、修正作業に着手する事ができること。
- (4) 音声の録音からテキストの編集、文字起こし内容の出力まで 1 つのアプリケーション内で行えること
- (5) 視聴用に音声認識で文字化されたテキストや文字起こし内容を別の画面に表示できること。
- (6) 音声の再生やテキストの編集は、キーボード操作だけでも行えること。
- (7) 音声認識エンジンを簡易カスタマイズできる単語登録機能が搭載されていること。
- (8) オフライン（インターネット未接続）の状態でも音声認識が可能なスタンドアローン型であること。
- (9) アプリケーションは、端末（Windows パソコン）にインストールして利用できること。
- (10) ライセンスの提供形態は、次のとおりとする。
 - ・決められた台数の端末にアプリケーションをインストールして利用できるスタンドアローンライセンス（ノードロック）。

2. 音声認識・編集機能

2-1. 音声認識機能

<共通機能>

- (1) 音声認識機能は、不特定話者対応で事前に話者の音声の登録・学習が不要であること。
- (2) 音声認識機能は、単語認識ではなく連続音声認識であり、発話内容を一字一句文字化できること。
- (3) 話し言葉（特に会議特有の話し言葉）が音声認識できること。
- (4) 音声認識の速度を調整できること。
- (5) ユーザー辞書に登録された単語を用いて音声認識ができること。
- (6) 音声認識を停止せずにユーザー辞書に単語を追加登録でき、かつ、即時反映され文字化できること。
- (7) 音声認識エンジンやユーザー辞書に登録されていない単語が認識、編集されたとき、未知語として表示され、単語登録ができること。
- (8) 自動的に句読点の出力ができること。
- (9) 話者振り分けができること。また、振り分けられた話者情報から自動で話者識別できること。

2-2. リアルタイム音声認識機能

- (1) 複数のサウンドデバイスを選択し、音声を入力できること。
- (2) 入力した音声をリアルタイムで音声認識できること。
- (3) 選択したサウンドデバイスごとに発言者の名前を割り付けることができ、その発言者ごとに認識結果を表示することができること。

2-3. 音声ファイル認識機能

- (1) 音声ファイル（MP3,WMA,WAV, M4A）、動画ファイル（WMV,MP4）に対し、音声認識できること。

3. 表示・編集機能

3-1. 表示機能

- (1) 認識結果を発話時間、発話者、発言内容の項目に分けて表示できること。
- (2) 発話時間の昇順で表示できること。

(3) 認識結果、編集結果を別の画面に表示できること。

3-2. 編集機能

(1) 発言内容単位で音声を聞きながら認識文字列を修正・編集できること。(バッチでの音声認識時)

(2) 発言内容の削除ができること。

(3) キーボードショートカットのカスタマイズが自由に設定できること。

(4) キーボードとマウスを使用して編集作業ができること。

(5) キーボードのみでも(マウスを使わなくても)編集作業ができること。

(6) 編集結果を Microsoft Word、Microsoft Excel、テキストファイルで出力できること。

(7) 音声データのみを音声ファイル(WAV または M4A)へ出力ができること。

(8) 重要な単語(キーワード)を設定する事で、テキスト化された発言内のキーワードをハイライト表示できること。

(9) 発言内容ごとに、「重要チェック」「決定事項」等のタグをつける事ができ、編集中やテキストファイル出力後に該当箇所を確認しやすくできること。

(10) 発言内容に対して絞り込み検索ができること。

(11) 絞り込み検索において、タグ、出席者、キーワード等で検索を絞り込むことができること。

(12) 絞り込み検索で絞り込んだ内容のみをテキストファイルに出力できること。

3-3. 音声再生機能

(1) 一つの発話の編集が終了したら、自動的に次の発話の編集に入り、音声も自動再生できること。(バッチでの音声認識時)

(2) ピッチを変えずに音声の再生速度の変更ができること。

(3) 選択した発言内容の音声をループ(繰り返し)再生できること。

3-4. 保存機能

(1) 音声認識した音声や認識結果、編集結果は自動的に保存できること。

(2) 保存したデータは一覧で表示できること。

(3) 保存したデータを削除できること。

(4) データ保存フォルダを任意の場所に変更できること。

4. 動作環境・ハードウェア仕様

4-1. 動作環境

下記スペックの端末で動作すること。

(1) プロセッサ：インテル® Core™ i5-8250U CPU @ 1.60GHz 1.80GHz

(2) メモリ：8GB

(3) ハードディスク：1GB 以上の空き容量 (インストール時)

(4) OS：Windows® 10 日本語版 64bit 版

5. インストール条件

(1) Web サイトから最新版のインストーラーをダウンロードできること。

(2) Microsoft Edge (EdgeHTML 版)、Microsoft Edge (Chromium 版)、Google Chrome (最新版)、Mozilla Firefox (最新版) に対応していること。

(3) 管理者権限でのインストールが可能なこと。

6. 保守サポート

(1) メールによるサポートサービスを行うこと。

(2) 製品仕様に関する問い合わせや問題発生時に問題解決支援を行うこと。

(3) 問合せ窓口はメールでの問い合わせ以外にも電話での問合せ窓口を用意すること。

- (4) 電話での問い合わせは平日（月～金：祝日除く）の9：00～17：30の間は問合せ可能な体制を構築すること。なお、専用窓口（即時回答可能な要員）を設ける必要はない。メールによる問い合わせについては、24時間365日受付可能とすること。
- (5) 不具合発生時に復旧対応を行うこと。
- (6) アプリケーションの定期バージョンアップを行うこと。
- (7) 製品バグフィックス及び最新マイナーバージョンアップモジュールの提供サービスを行うこと。
- (8) WindowsのOSやWebブラウザのバージョンアップに対応すること。