

## AI 議事録作成支援システムライセンス調達仕様書

### 1 事業目的

A I 等の先端技術を導入し、音声を自動的に認識して文字化（以下「テキスト化」という。）を行うシステムを導入することで、庁内の会議・審議会等（以下「会議等」という。）の議事録の作成に要する職員の作業時間を削減し、業務の効率化と市民サービスの向上を図る。

### 2 調達概要

- (1) 会議録作成のため会議録支援システムをオンプレ版にて提供すること。
- (2) 音声認識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 調達内容および調達要件

音声認識エンジンは、株式会社アドバンスト・メディア「AmiVoice」もしくは、AmiVoice と同等以上の変換精度を有するものとし、別紙『音声認識システム仕様書』のとおりとする。

なお、AmiVoice 以外の音声認識エンジンを調達する場合は、同等以上の変換精度を有する証明を行い、浦添市の承認を得ること。また、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

令和8年6月1日から令和10年3月31日まで

### 5 納入場所及び納入期限

納入場所) 浦添市役所 企画部情報政策課

納入期限) 令和8年5月31日迄

### 6 納品物

音声認識システム『AmiVoice ScribeAssist』4ライセンス

### 7 見積（入札）金額

- (1) 令和8年6月1日～令和10年3月31日までの使用期間（22ヵ月分）における総額料金  
※4ライセンス料を含む  
※初期費用または基本料金等が発生する場合は、見積（入札）額に含めた額とすること。  
※「AmiVoice」以外の音声認識エンジンの場合は、あらゆる諸費用込みの定額月額とする総額とする。  
※事前に見積内訳書を提出すること。

### 8 請求及び支払い方法

- (1) 使用料は、月額料金とする。
- (2) 使用料は、毎月後払いとする。

### 9 その他

- (1) 導入機器の円滑な運用のために委託者を支援するとともに、委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委託者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。

(5) 代理店証明できる書類の提出を行うこと。

10 特記事項

- (1) 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。
- (2) 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、委託者と協議の上決定する。

## 音声認識システム仕様書

### 1. システム仕様

#### 1-1. 基本機能

- (1) 会議の開催形式（対面または Web 会議などの非対面、さらにはそのミックスとなる形式）、会議の種類（地方議会、団体内での社内会議や面談、外部顧客や取引先との商談、株主総会、記者会見・取材）、および人数の多寡などにかかわらず幅広いシーンで利用できるスタンドアロ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) 領域特化型（議会、医療・製薬、建設）と英語の音声認識辞書に対応していること。
- (4) 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ことができること。
- (5) 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。出力先は、Word や Excel、CSV（カンマ区切り）、テキストファイルから選択できること。
- (6) 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（字幕ポップアップ機能）。かつ、字幕のフォント、サイズ、背景色、透明度の変更が可能であり、ルビも付けられること。  
Web 会議サービス「Zoom」向けには本機能とは異なり、API による字幕送信が直接可能であり、Zoom 画面上に字幕の表示が可能であること。
- (7) 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。
- (8) 音声認識辞書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。
- (9) 別途契約により音声認識辞書を個別にカスタマイズできること。
- (10) 全文を起こす議事録作成に適したシンプルなモード（エディットモード）、会議主催者の進行補助機能としてアジェンダ設定やタイムキーパー機能、自由に議事メモを挿入できるエリアを具備したモード（ファシリテーションモード）の両方が備えられており用途に合わせて使い分けができること。
- (11) 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ン型であること。
- (12) 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。
- (13) ライセンスの提供形態は、以下の 4 種類から選択できること。
  - ・インストール台数に制限がなく、外部のクラウド上のライセンス管理サーバーを用いて、ライセンス数の範囲内でアプリケーションが同時に利用できるクラウドサーバーライセンス。
  - ・インストール台数に制限がなく、内部のネットワーク内に構築したライセンス管理サーバーを用いて、ライセンス数の範囲内でアプリケーションが同時に利用できるオンプレミスサーバーライセンス。
  - ・インストール台数に制限がなく、端末に USB キーを接続した時にのみアプリケーションが利用できるスタンドアロンライセンス（USB キー）。
  - ・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ンライセンス（ノードロック）。
- (14) インターネットに接続が必要な機能は、プロキシ経由であっても利用できること。
- (15) 文字化されたテキストや文字起こし内容を要約する AI 要約機能が搭載されていること。要約の方法として、ScribeAssist アプリから要約ができるモード（ScribeAssist で要約）、クラウド上に音声データを送信し、「議事録エディタ」という専用の編集環境で要約および編集を行うモード（議事録エディタで要約）、すべての処理を PC 内部（ローカル環境）で完結する自動要約（スタンドアロン要約）が備えられており、用途に合わせて使い分けができること。
- (16) クラウドの AI 要約機能のテキストは、Microsoft 社の Azure OpenAI Service を利用して生成できること。Azure OpenAI Service に通常含まれる「データの 30 日間保持」および「同社の従業員による有害コンテンツ・不正使用の監視」のいずれも対象外であること。
- (17) クラウドの AI 要約機能では、入出力するテキストの最小単位をトークンとし、処理するテキスト量に

応じてトークン数を計測できること。1 ライセンスの契約あたり、月々最大 300 万トークンまで要約機能を利用できること。

- (18) クラウドの AI 要約機能は、ユーザー認証 (VoXT One アカウント) により利用でき、そのアカウント管理や当月の残トークン数を確認できる管理者サイト (VoXT One 管理者サイト) を提供すること。また、ユーザーをまとめるグループ単位で、要約機能の利用可否を設定できること。
- (19) 多様な色覚の方が不自由なく利用できるよう、カラーユニバーサルデザイン認証を取得していること。

## 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) 話し言葉の音声認識ができること。
- (4) 領域特化型(議会、医療・製薬、建設)と英語の音声認識辞書に対応していること。
- (5) 音声認識が出来ない固有名詞に対しては表記と読みを付与することで音声認識が可能になる単語登録機能を持つこと。単語は複数のグループで管理ができ、シーンに応じて適用する単語グループを選択できること。
- (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) 音声認識辞書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、未登録単語として表示され、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) AI による自動話者振り分け機能があること。また、振り分けられた話者情報から自動で話者識別ができること。
- (10) 話者識別機能は学習データのインポート、エクスポート機能を有すること。

### 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ができること。
- (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。
- (4) 音声認識中であってもデバイス設定を変更できること。

### 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、音声認識ができること。
- (2) 複数の音声ファイルを指定できること。(最大 10 件)
- (3) 音声ファイルの全てに対して音声認識を行うのか、指定した時間の範囲に対して音声認識を行うか、何れかの選択ができること。
- (4) 音声ファイル全体の認識が全部終わっていない状態でも、認識処理が完了した部分から音声再生を含めた編集の開始が可能であること。

## 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) 音声を波形で表示できること。
- (4) 発話の開始時間の変更ができること。

### 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) 編集結果の項目を選択して Microsoft Word、Microsoft Excel、CSV(カンマ区切り)、テキストファイルに出力できること。Microsoft Word の場合は、指定した書式で出力できること。
- (7) 音声データのみを音声ファイル(WAV または M4A)に出力できること。
- (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「検索・置換」または「絞り込み検索」ができること。
- (11) 検索・置換において、検索結果を 1 件ごとに置換、またはすべて置換ができること
- (12) 絞り込み検索において、タグ、出席者、キーワード、表示列の文字数等で検索を絞り込むことができること。
- (13) 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。
- (14) 発言リストに対して、編集操作の取り消し (Undo) およびやり直し (Redo) が行えること。
- (15) 複数の発言を同時に選択でき、選択した発言に対して一括で削除、タグ付け、発言者の設定が行えること。

### 3-3. 音声再生機能

- (1) 一つの発言の編集が終了したら、自動的に次の発言の編集に入り、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。

### 3-5. 外部サービス連携機能

- (1) Zoom に音声認識結果を字幕として表示できること。

## 4. クラウド要約機能

### 4-1. 基本機能

- (1) クラウドとはインターネット回線で通信を行い、TLS1.2/TLS1.3 (暗号強度 256 ビット) で暗号化ができること。
- (2) ユーザーのアカウントを設定、認証に成功することで、クラウドの AI 要約機能を有効化できること。
- (3) ユーザーのアカウントを管理する管理者サイトを提供すること。アカウントは、1,000 件まで登録できること。

### 4-2. ScribeAssist で要約

- (1) 認識結果や編集結果の日本語 (または英語) の文字情報をクラウドに送信し、自動的に要約された結果を取得できること。
- (2) 要約の方式として、以下の 2 種類から選択できること。

- ・すべて AI にまかせる：文字起こし結果から自動的に議題の切り替わりを探し、その議題ごとに議題名と要約を生成
  - ・設定した議題ごとに要約する：予め設定した議題ごとに要約を生成
- (3) 送信した文字情報、要約結果の文字情報は、クラウド上に残さないこと。
- (4) 要約結果の文字情報は、クリップボードにコピーできること。また、Microsoft Word に出力できること。

#### 4-3. 議事録エディタで要約

- (1) 音声認識した音声および認識結果、編集結果の文字情報をクラウドにアップロードした後、ブラウザ上の編集環境（議事録エディタ）で要約や編集が行えること。
- (2) 要約の出力形式として、以下の2種類から選択できること。
- ・議事録：議題ごとに概要・決定・ネクストアクションを出力
  - ・要点まとめ：議題ごとに要点のまとめを出力
  - ・カスタム：自由にプロンプトを書いて出力
- (3) カスタム用のテンプレートを複数登録できること。また、要約の際にそれを選択できること。
- (4) 出力形式が「議事録」、「要点まとめ」の場合は、要約の方式として、以下の2種類から選択できること。
- ・すべて AI にまかせる：文字起こし結果から自動的に議題の切り替わりを探し、その議題ごとに議題名と要約を生成
  - ・設定した議題ごとに要約する：予め設定した議題ごとに要約を生成
- (5) 出力形式が「議事録」、「要点まとめ」の場合は、要約の根拠を可視化するために、元のデータのどの部分を要約したのかを表示できること。
- (6) 要約結果を URL リンクで他者と共有ができること。その際、パスワードと有効期限を設定できること。

#### 4-4. 管理者サイト (VoXT One 管理者サイト)

- (1) 管理者のアカウントを管理（登録・修正・削除）できること。
- (2) ユーザーのアカウントを管理（登録・修正・削除）できること。
- (3) グループの概念を有し、ユーザーのグルーピングができること。
- (4) グループごとにクラウドの AI 要約機能の利用可否を設定できること。
- (5) 議事録エディタで要約、編集したデータの削除期限を設定できること。
- (6) アクセス履歴を確認できること。その内容を CSV ファイルにエクスポートできること。
- (7) トークン利用履歴を確認できること。その内容を CSV ファイルにエクスポートできること。
- (8) 当月に使用できる残トークン数を確認できること。
- (9) アカウントに対し、定期的にパスワードを変更させるように日数を設定できること。
- (10) ユーザーサイトや管理者サイトへのアクセスを許可する接続元の IP アドレスを設定できること。設定できる IP アドレスは、1,000 件まで登録ができること。

#### 5. スタンドアローン要約

- (1) すべての処理を PC 内部（ローカル環境）で完結し、インターネット接続を必要とせずに生成 AI による自動要約を利用できること。
- (2) 機密性の高い情報も外部に送信せず、安全に処理できること。
- (3) 利用量に制限なく使用できること。
- (4) 要約する際のオプションとして、以下の3種類の出力形式を選択できること。
- ・議事録形式：議題ごとに概要・決定事項・ネクストアクションを出力
  - ・要点まとめ形式：議題ごとに要点をまとめて出力
  - ・カスタム形式：ユーザーが定義したフォーマットで出力
- (5) 要約結果を並べて比較でき、プロンプトの改善が容易に行えること。

- (6) Lite 版（スピード重視のみ利用可）及び Full 版（「スピード重視」または「精度重視」を利用可）の選択ができること。
- (7) 発言者氏名を含めた要約ができること。
- (8) 設定した議題ごとに要約できること。

## 6. 動作環境・ハードウェア仕様

### 6-1. 動作環境

- (1) CPU：インテル® Core™ i5 プロセッサ（ターボ・ブースト利用時の最大周波数 2.5GHz）以上
- (2) メモリ：8GB 以上
- (3) ハードディスク：1GB 以上の空き容量（インストール時）
- (4) CD-ROM ドライブ：インストール時に必要
- (5) OS：Microsoft Windows® 10 / Microsoft Windows® 11
  - ※日本語版、64 ビット版(x64)に対応。32bit 版(x86)には対応していません。
  - ※Mac、Linux には対応していません。
  - ※.NET Framework4.6.0 以降がインストールされていること。
  - ※仮想化環境（仮想デスクトップ、仮想アプリケーション）ではインストールや動作を保証いたしかねます。
  - ※移動ユーザープロファイル環境ではインストールや動作を保証いたしかねます。
- (6) その他ハードウェア：マイク入力端子・ヘッドフォン端子（または USB ポート）
  - ※音声入力や音声再生時に必要となります。

### 6-2. 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）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。

### 6-3. 議事録エディタ

- (1) OS：Windows 11 Pro
- (2) Web ブラウザ：Microsoft Edge、Google Chrome
  - ※上記以外の OS や仮想環境（VDI や仮想ブラウザ等）、リモートデスクトップ接続では正しく表示されない、または正しく機能しない場合があります。
  - ※上記以外の Web ブラウザでは、アクセスすることができません。
  - ※外部アクセスへの URL フィルタリング、マルウェア対策、通信の復号化対応等によりご利用できないことがあります。
  - この場合は、以下の URL をホワイトリストにご登録ください。ただし、LGWAN から接続して利用する場合は、登録不要。
    - ・ wss://pro.vox.jp
  - ※wss:WebSocket Secure の略。議事録エディタがサーバーとの通信で使用する通信プロトコルです。

## 7. 保守サポート

- (1) Web 問い合わせフォームによるサポートサービスを行うこと。
- (2) 不具合発生時に復旧対応を行うこと。
- (3) アプリケーションのバージョンアップを行うこと。
- (4) 製品バグフィックス及び最新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。
- (5) 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。