

## 議事録作成支援システム要件

### 1 議事録作成支援システム

#### (1) 基本機能

- ① 会議のように複数の発言者による自然発話の音声を自動的に文字化する音声認識システムをクラウドサービスで提供できること。
- ② クラウドサービスの提供のため、インターネット上にサイト（以下、サービスサイト）を公開すること。サービスサイトに音声ファイルや動画ファイルをアップロードすること（以下、認識依頼）により、音声認識システムが自動的に文字化を行い、それにより作成されたファイル（以下、認識結果ファイル）をサービスサイトからダウンロードできること。
- ③ 音声ファイルは、WAV, WMA, MP3 及び M4A の全ての形式に対応していること。
- ④ 動画ファイルは、WMV, AVI, MP4 及び M4V の全ての形式に対応していること。
- ⑤ 認識依頼してから認識結果ファイルのダウンロードが可能になる状態でまで（以下、処理完了）の時間が、音声ファイルの長さと同等の時間内であること。
- ⑥ 認識結果編集ソフトを用いて、正しく文字化されなかった文字を修正したファイル（以下、修正ファイル）をサービスサイトにアップロードすることで、音声認識システムの文字化の精度向上を図れること。
- ⑦ 認識依頼時や処理完了時に通知メールをユーザーごとに、事前に設定したメールアドレスに自動送信ができること。
- ⑧ 修正ファイルをアップロードできること。
- ⑨ ユーザーID について、100 ユーザー以上発行できること。

#### (2) セキュリティ要件

- ① インターネット上の通信は、TLS1.2 以上（暗号強度 256 ビット）で暗号化ができること。
- ② ユーザーID とパスワードによるユーザー認証ができること。
- ③ ユーザーID ごとに次の二種類の何れかの権限を設定できること。
  - ・ 一般ユーザー
  - ・ 管理者ユーザー
- ④ 障害発生時は、メール通知または Web ページ等に掲載すること。
- ⑤ アップロードした音声ファイルは、一定期間で削除されること。
- ⑥ 処理依頼に係る音声ファイル及び認識結果ファイルへのアクセスは、当該処理依頼を行ったユーザーのみが行えること。

### (3) 管理者機能

- ① ユーザーの追加・更新・削除等ができること。
- ② 全てのユーザーが行った認識依頼に係る音声ファイル及び認識結果ファイルへアクセスできること。
- ③ 全てのユーザーの利用状況を確認できること。
- ④ 全てのユーザーの操作ログ及びアクセスログを CSV 形式で取得できること。管理者機能で対応できない場合は、市の要求に応じて提出すること。

### (4) 動作環境

- ① 下記 OS に対応していること。
  - ・ Windows 10 Enterprise LTSC 64bit
  - ・ Windows 10 Enterprise LTSC 64bit
  - ・ Windows 10 Professional 64bit
  - ・ Windows 8.1 Professional 64bit
- ② 下記 Web ブラウザに対応していること。
  - ・ Microsoft Edge (Windows10 版)
  - ・ Google Chrome (最新バージョン)

## 2 認識結果編集ソフトウェア

### (1) 表示機能

- ① 発話者、発話開始時間、発話終了時間、認識結果文字列、編集結果文字列、編集者名、備考が並べて表示できること。その際、必要な項目列のみを表示できること。
- ② 発話区間の開始時刻と終了時刻が画面上で確認できること。
- ③ 候補単語の表示は、文字列の直下もしくは編集枠右側の専用枠、もしくは非表示へ切り替えができること。
- ④ 音声信号は波形表示され、発話区間も視覚的に表示できること。
- ⑤ 音声波形表示の時間軸及びレベルを縮小・拡大できること。
- ⑥ 波形表示位置を移動させるためのスクロールバーがあること。その際、全体に対する表示枠の比率が、スライド枠の長さに対する比率になるようにスクロールバーの長さが変化すること。

### (2) 編集機能

- ① 発話区間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（以下、編集モード機能）。
- ② 発話区間単位でなく、議事録に近い体裁でも音声を聞きながら認識文字列を修正・編集できること（以下、校正モード機能）。
- ③ 選択した文字のかな漢字再変換（IME 機能）ができること。
- ④ 発話区間以外の区間の音声は自動的にスキップできること。
- ⑤ キーボードショートカット及びツールバーのカスタマイズが自由に設定できること。

- ⑥ キーボードとマウスを使用して編集作業ができること。また、マウスを使用することなく、キーボードのみでも編集作業ができること。
- ⑦ 改行マーク及び字下げの場合などのスペース入力ができること。
- ⑧ 改行を含む複数行表示・編集ができること。
- ⑨ 行間の間隔を変更できること。
- ⑩ 文字のフォントやサイズを変更できること。
- ⑪ 折り返しの文字数を変更できること。
- ⑫ 認識結果及び編集結果の文字列に対して全文検索ができること。検索結果の発言音声即座に再生できること。
- ⑬ 認識結果及び編集結果を Microsoft Word 形式又はテキストファイル形式で出力できること。
- ⑭ 特定のキーに対し、指定の単語を 10 個以上割り当てられること。
- ⑮ 予め設定された指定の単語、文章を選択し挿入できること。
- ⑯ 前面に表示されるメモ領域があり様々な用途に活用できること。(例：書き起こしルールの記載、よく使う文章等の記載)
- ⑰ 認識結果及び編集結果の文字列や音声にインデックス（マーカー）を付けることができ、編集時に目的の位置を素早く見つけたり、検索したりすることができること。

### (3) 編集モード機能

- ① 編集ビュー直下に次の発話の認識文字列や編集結果文字列を表示できること。
- ② 単語単位で認識文字列の複数の候補が選択修正できること。
- ③ 発話区間の時間位置の変更ができること。
- ④ 発話区間の作成及び削除ができること。
- ⑤ 発話区間の結合と分割ができること。
- ⑥ 現在の音声再生の開始位置に該当する発話がショートカットキーで選択できること。
- ⑦ 編集結果として、発話者、発話開始時間、発話終了時間、認識結果文字列、編集結果文字列の CSV ファイルが出力できること。その際、出力した CSV ファイルに各項目の見出しを付加できること。
- ⑧ 修正後のデータをファイルとして書き出す場合、自動的に編集者別に書き出しができること。
- ⑨ 分割・結合処理の場合、各人が修正した部分について、自動的に編集者の表示が出力されること。
- ⑩ 事前に登録した名前を用いて、話者名を編集できること。
- ⑪ 話者名を一括編集できること。
- ⑫ 話者名を一括置換できること。

### (4) 校正モード機能

- ① 行番号を表示したり非表示にしたりして、修正できること。
- ② 現在の音声再生の開始位置に該当する認識文字列や編集結果文字列がショートカ

ットキーで選択できること。

- ③ 認識文字列や編集結果文字列のカーソル位置に該当する音声開始位置がショートカットキーで選択できること。

#### (5) 音声再生機能

- ① 一つの発話の編集終了後、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。
- ② 複数の音声チャンネルを同時もしくは個別に聞けること。個別に音声レベルの調整ができること。
- ③ 複数の音声チャンネルが存在する場合、選択した発話のチャンネル以外の音声が小さくなること。それにより該当音声聞きやすくなること。
- ④ 認識結果の文字列に対するカーソル位置と音声再生の開始位置がリンクしていること。
- ⑤ 選択発話の0～5秒程度前から音声再生でき、選択発話の0～5秒程度後まで音声再生できること。秒数はユーザー側で任意に設定ができること。
- ⑥ 音声の再生、停止、早戻し、早送り、発話区間単位の前後移動がキーボード操作、マウス操作いずれでも簡単にできること。
- ⑦ 画面が他のソフトに隠れてもショートカットにより再生操作ができること。
- ⑧ ピッチを変えずに音声の再生速度の変更ができること。
- ⑨ 発話区間の印のバーを選択すると、その音声再生できること。
- ⑩ 選択した発話区間の音声をループ（繰り返し）再生できること。
- ⑪ 音声の再生位置に合わせて認識結果編集カーソルを自動移動できること。
- ⑫ 音質を調整して音声が聞きやすくなること。
- ⑬ 自動で一定の音量に調整する機能を有すること。
- ⑭ マーカーが付けられた音声をマスク（ビープ音等に置換）して、再生できること。

#### (6) その他機能

- ① **USB** キーを用いてソフトウェアの起動制限ができること。
- ② **USB** キーでのソフトウェアの起動制限の他にライセンスファイルでの起動制限（有効期限による制限）ができること。
- ③ **USB** キーでのソフトウェアの起動制限の他にライセンスサーバに接続して起動制限（有効期限・同時接続数による制限）ができること。
- ④ 音声ファイル（WAV, WMA, MP3, WMV）から音声認識処理を介さず、認識結果のテキストがなくとも、編集できること。
- ⑤ 編集用のファイルの任意の選択箇所を抜き出し、ファイルに保存できること。
- ⑥ 編集用のファイルから音声データのみを音声ファイルへ書き出しができること。
- ⑦ 音声データの任意の選択箇所を抜き出し、音声ファイルへ書き出しができること。
- ⑧ マーカーが付けられた音声データをマスクして、音声ファイルへ書き出しができること。
- ⑨ 任意の音声データで発話の自動検出ができること。発話区間以外の区間の音声は自

動的にスキップできること。その際、発話の最大発話長を設定できること。

- ⑩ ライセンスキーの追加ができること。(他で作成した編集用のファイルもライセンスキーを追加する事で使用可能とできる)
- ⑪ 編集用のファイルは、ライセンス証書に記述された ID 以外では取り扱うことができないこと。
- ⑫ チャンネル別に音声のみのファイルに保存して任意の音声プレイヤーで再生できること。
- ⑬ 複数の編集用のファイルに対して条件付全文検索ができること。
- ⑭ 編集用のファイルの未読・既読・編集中の状態がアイコン表示で判断できること。
- ⑮ 編集データに対する仕分けや複数の編集用のファイルの編集結果をひとつのファイルに書き出す等一括処理ができること。
- ⑯ カスタマイズされたテンプレートを使用して指定の書式の Word ML ファイルへ出力できること。
- ⑰ マーカーが付けられた認識文字列や編集結果文字列の書式(文字飾り)を変更して、Word ファイルや一太郎ファイルへ書き出しができること。
- ⑱ 編集用のファイルの開始時間の指定・修正ができること。
- ⑲ 前回保存した時点での作業状態を保持し、次回ファイルを開いた時、前回の作業状態から編集が開始できるレジューム機能を有すること。

#### (7) 動作環境

- ① 下記 OS に対応していること。
  - ・ Windows 10 Enterprise LTSC 64bit
  - ・ Windows 10 Enterprise LTSC 64bit
  - ・ Windows 10 Professional 64bit
  - ・ Windows 8.1 Professional 64bit
- ② 下記 Web ブラウザに対応していること。
  - ・ Microsoft Edge (Windows10 版)
  - ・ Google Chrome (最新バージョン)
- ③ ネットワーク環境に依存せず動作すること。