

AI音声認識文字起こしサービスに関する情報提供依頼書（RFI）_質問回答表

項番	質問	回答
1	アップロードされた音声・動画データおよび文字起こし結果データの保存期間について、要件・想定はありますでしょうか(要件2-10では利用状況情報の1年以上の保存が示されていますが、データ本体の保存期間の記載がないためお伺いします)。	音声・動画データ及び文字起こし結果データの保存期間の要件については、「1年以上」とする見込みです。 推奨の期間がございましたら、ご提示ください。
2	外国語対応について、例示されている英語・中国語・韓国語のほかに、優先的に対応が必要な言語があればご教示ください(窓口での外国籍の方とのコミュニケーションで利用頻度の高い言語 等)。	現時点で外国語対応に関する要件は未確定です。 対応可能な言語について、すべてご提示ください。
3	「提供するクラウド環境は、政府情報システムのためのセキュリティ評価制度 (ISMAP) 又はISMSクラウドセキュリティ認証を取得しているサービスであること」について、サービスの基盤となるクラウド基盤 (IaaS/PaaS) がISMAP登録済みであることをもって本要件を満たすと解してよいでしょうか。 それともSaaSサービス自体のISMAP登録またはISMSクラウドセキュリティ認証の取得が必要でしょうか。	SaaSサービス自体のISMAP登録もしくはISMSクラウドセキュリティ認証を取得していただく必要があります。
4	貴市専用のサービスとして構築・提供する形態を想定しているため、デジタルマーケットプレイスへの登録は予定しておりません。 登録のないサービスであっても、情報提供および令和9年度の調達への参加は可能という理解でよろしいでしょうか。	現時点でデジタルマーケットプレイスへの登録は必須要件ではございませんが、調達時にデジタルマーケットプレイスを利用して調達を行う可能性もあるため、デジタルマーケットプレイスへの登録をお願いいたします。
5	仮想ブラウザ (SASTIK) 経由で本サービスへ接続する場合の送信元 (出口) グローバルIPアドレスについて、固定IPアドレスか、また貴市専用か他団体との共用かをご教示ください。 接続元IPアドレスによるアクセス制限 (要件6-28) をSASTIK経由の利用にも適用できるかを検討するためです。	長野市専用で利用する固定IPアドレスになります。
6	利用経路に関する要件1-4 (LGWAN-ASP)、1-5 (インターネット接続系)、1-6 (仮想ブラウザSASTIK) については、いずれかに該当する形であれば問題ないという理解でよろしいでしょうか(確認)。 あわせて、本サービスを利用する職員の主な端末環境 (LGWAN接続系・インターネット接続系の構成比等) をご教示ください。	利用経路の要件は、LGWAN-ASP・インターネット接続系・仮想ブラウザの3つの経路全てにおいて利用可能であること、とする見込みです。 また、LGWAN接続系端末は、ほぼ職員1人に対し1台の利用となっていますが、インターネット接続系端末は所属ごとに数台配備の上、職員にて数台の端末を共用しています。端末構成比はLGWAN接続系とインターネット接続系とで、概ね5対1になります。
7	仮想ブラウザ (SASTIK) 経由での利用について、想定される操作範囲をご教示ください(文字起こし結果の閲覧・ダウンロードまでを想定か、音声・動画ファイルのアップロードやマイク入力によるリアルタイム文字起こしの利用も想定か)。 また、SASTIK環境での動作検証の機会 (トライアル時の検証環境のご提供等) はありますでしょうか。	仮想ブラウザ (SASTIK) 経由での利用については、文字起こし結果の閲覧及びダウンロードを行うことを想定しており、音声・動画ファイルのアップロードやマイク入力によるリアルタイム文字起こしでの利用は想定しておりません。 また、当市のSASTIK環境には検証環境はありませんが、ご要望に応じ事前に本番のSASTIK環境を用いた動作検証の機会を設けることは可能です。

AI音声認識文字起こしサービスに関する情報提供依頼書（RFI）_質問回答表

項番	質問	回答
8	窓口での対面コミュニケーション支援について、想定される設置窓口数(箇所数)をご教示ください。 また、利用に必要な機器(音声入力用のマイク、サービスを動作させる端末、表示用のスクリーン・ディスプレイ・タブレット等)は貴市にてご用意される想定でしょうか、それとも機器の提供・設置も本サービスの提供範囲に含める想定でしょうか。	窓口での対面コミュニケーション支援において、現時点で想定する設置窓口数は1か所になります。 また、サービス利用に必要な各種機器も提供範囲に含まれます。
9	動画データの文字起こしについて、主な想定用途をご教示ください(例: Web会議(Teams/Zoom等)の録画データからの議事録作成、議会・委員会等の中継録画の文字起こし、広報・研修動画への字幕付与 等)。 あわせて、音声・動画を合わせた月間の想定処理時間(文字起こし対象の総時間)の目安があればご教示ください。	基本的には記載いただいたような、会議等の動画について文字起こしを行うことでの利用を想定しています。 処理時間の想定はありませんので、上限時間についてご提示ください。
10	概算費用算出の前提とするため、想定される利用規模をご教示ください(利用部署数、利用者数またはアカウント数、リアルタイム文字起こしの同時利用数の目安 等)。	所属数150、職員数4,000が利用する前提としてください。なお、職員には個人のメールアドレスは付与されておりませんので、アカウントにメールアドレスを使用する場合は、所属のアカウントを同じ所属の職員で共有して利用する方法を想定しています。 また、リアルタイム文字起こしの同時利用数については現時点で未確定となっております。同時利用数の上限についてご提示ください。