



Aplikasi Voice To Text Sekretariat DPRD Kabupaten Hulu Sungai Tengah

Muhammad Abdul Ghani¹, Abdul Hamid², Cucut Annaningtyas³

^{1,3}Manajemen Informatika, Politeknik Harapan Bangsa Surakarta, Surakarta, Indonesia

²Sistem Informasi, Institut Teknologi Saptamandiri, Balangan, Indonesia

E-mail: ghaniriaelm@gmail.com , hamied@itsmandiri.ac.id , valerangga2006@gmail.com

*Corresponding Author

Article History: Received: July 31, 2023; Accepted: August 30, 2023; Published: December 31, 2023

ABSTRACT

Meetings are routine events in the function of legislation. The main purpose of the meeting is to record the minutes of the meeting accurately and completely, which is time consuming and carries the risk of typos. Therefore, this study aims to implement a voice to text application at the DPRD Secretariat of Hulu Sungai Tengah Regency in order to improve the efficiency of recording the minutes of meetings. This application uses speech recognition technology to automatically convert the speech of meeting participants into text. This research examines several meetings and measures the time required to enter meeting notes manually and use a voice to text application. The results show that using this application can save up to 50% time and reduce the risk of typos. Apart from that, this application allows users to edit and correct text easily. Therefore, this study recommends the implementation of the voice to text application at the DPRD Secretariat of Hulu Sungai Tengah Regency as an effective alternative to increase the efficiency of recording the minutes of meetings.

Keywords: *voice to text application, Secretariat of DPRD Hulu Sungai Tengah Regency, efficiency, minutes of meetings, speech recognition technology*



Copyright © 2023 The Author(s)

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

PENDAHULUAN

Rapat merupakan kegiatan yang rutin dilakukan dalam fungsi legislasi. Salah satu tujuan utama dalam rapat adalah menyusun risalah rapat yang akurat dan lengkap. Risalah rapat adalah kegiatan mencatat atau menyalin seluruh hasil dari pertemuan (Darwanto, Saputra, and Wardana 2021), namun pekerjaan ini seringkali membutuhkan banyak waktu dan berisiko terjadi kesalahan penulisan. Selain itu, tugas ini juga membutuhkan ketelitian dan ketekunan.

Voice to text adalah teknologi yang merubah suara dari mikrofon menjadi teks di layar komputer. Pengenalan ucapan adalah istilah dalam ilmu komputer, dikenal sebagai pengenalan ucapan otomatis, mengubah audio menjadi teks. Keuntungannya termasuk mengurangi kesalahan ketik. (Setia Buana, 2020)

Aplikasi *voice to text* telah menjadi salah satu inovasi yang signifikan dalam bidang *voice recognition* (Smith et al., 2022). Dalam jurnal yang dikutip, penulis menjelaskan bahwa aplikasi tersebut berpotensi mengubah cara kita berinteraksi dengan teknologi dengan mengubah ucapan menjadi teks yang dapat dibaca dan dipahami oleh mesin. Aplikasi ucapan-ke-teks menawarkan solusi untuk menghemat waktu dan tenaga serta meningkatkan akses ke informasi.

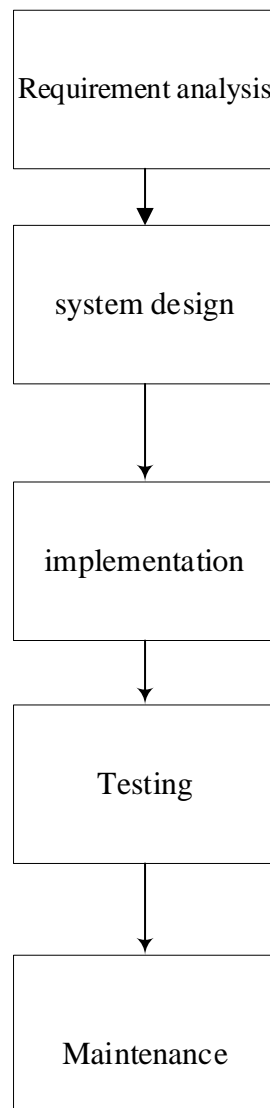
Mereka telah berhasil diterapkan di berbagai bidang seperti transkripsi wawancara, pemrosesan data audio, interaksi manusia-komputer (Smith et al., 2022). Inilah sebabnya mengapa penting untuk mengetahui lebih banyak tentang aplikasi ucapan-ke-teks dan potensinya di lingkungan modern untuk memaksimalkan kegunaannya dalam kehidupan sehari-hari.

Untuk mengatasi masalah waktu kerja yang menyita waktu dan resiko kesalahan dalam penyusunan notulensi rapat, penelitian ini mengusulkan solusi alternatif yang efektif yaitu penggunaan aplikasi *speech-to-text* dengan metode *speech recognition*. Aplikasi ini mengandalkan teknologi pengenalan ucapan untuk secara otomatis mengonversi ucapan peserta rapat menjadi teks. Dengan menggunakan teknologi ini, diharapkan dapat menghemat waktu pengetikan hingga 50% menit dan mengurangi risiko kesalahan penulisan. Kajian ini menunjukkan beberapa keterbatasan permasalahan terkait penerapan *speech-to-text* Sekretariat DPRD Kabupaten Hulu Sungai Tengah. Aplikasi ini hanya mendukung Bahasa Indonesia, tidak tersedia untuk bahasa lain. Kemampuan aplikasi untuk mengenali ucapan terbatas pada kecepatan suara normal, dan menghadapi tantangan pada kecepatan suara yang berbeda. Berbagai bahasa lisan dapat dikenali menggunakan metode pengenalan ucapan, tetapi Anda harus mengetahui aksen atau dialek yang dapat memengaruhi akurasi transkripsi. Aplikasi ini dirancang menggunakan bahasa pemrograman *Python*, sehingga membatasi kemungkinan implementasi dalam bahasa pemrograman lain. Khusus untuk rapat-rapat Kelurahan dan tidak terkait dengan rapat-rapat lainnya. Aplikasi ini hanya satu arah, untuk pidato gubernur distrik dan laporan hasil badan anggaran, aplikasi ini tidak dapat menjalankan fungsi *text-to-speech*. Walaupun dengan keterbatasan, diharapkan penelitian ini tetap memberikan solusi yang efektif untuk membantu pencatatan risalah rapat di Sekretariat DPRD Kabupaten Hulu Sungai Tengah. Pada penelitian ini, penulis akan mengevaluasi keefektifan aplikasi *speech-to-text* menggunakan metode *speech recognition* dalam meningkatkan efisiensi pencatatan risalah rapat dan sejauh mana aplikasi tersebut dapat menghasilkan risalah rapat yang akurat dan efisien di Sekretariat DPRD Kabupaten Hulu Sungai Tengah dalam Bahasa Indonesia. Penelitian ini akan dibatasi pada implementasi aplikasi *speech-to-text* pada rapat DPRD, menggunakan bahasa Indonesia, mengenali ucapan dengan kecepatan suara normal, dan membangun aplikasi menggunakan bahasa pemrograman *Python*. Diharapkan aplikasi ini menjadi sarana yang efektif dan efisien untuk mendukung alur kerja Sekretariat DPRD Kabupaten Hulu Sungai Tengah.

METODE

Penulis menggunakan pendekatan metode *voice recognition* dalam mengembangkan aplikasi perekaman dan transkripsi suara di lingkungan kantor. *Voice recognition* adalah teknologi yang memungkinkan perangkat untuk menangkap kata-kata yang diucapkan manusia ke dalam mikrofon. Kata-kata ini diproses oleh pengenalan ucapan dan akhirnya sistem mengeluarkan kata-kata yang dikenali (Vashisht, Pandey, and Yadav 2021). Metode ini memanfaatkan teknologi pengenalan suara untuk mengubah ucapan menjadi teks secara otomatis. Proses pengembangan aplikasi ini melibatkan identifikasi kebutuhan pengguna, perancangan sistem, implementasi teknologi *voice recognition*, pengujian kinerja, dan penyesuaian berdasarkan umpan balik pengguna. Proses-proses tersebut dijalankan secara berurutan dan terstruktur untuk mencapai hasil akhir yang sesuai dengan tujuan pengembangan aplikasi.

Penulis memanfaatkan metode pengembangan sistem menggunakan SDLC (*Software Development Life Cycle*) sebagai panduan dalam proses pengembangan aplikasi *voice recognition*. Studi literatur dilakukan untuk mendapatkan pemahaman tentang konsep dan teknologi terbaru yang relevan. Selanjutnya, dilakukan analisis kebutuhan pengguna, perancangan sistem, pengujian, dan evaluasi guna menciptakan aplikasi yang efisien dan sesuai dengan tujuan yang ditetapkan.



Gambar 1 Metode pengembangan sistem menggunakan SDLC

Pada tahap analisis kebutuhan penelitian Sistem Pelayanan Sekretariat DPRD Hulu Sungai Tengah, langkah pertama adalah memahami sistem pelayanan dan teknologi pengenalan suara yang digunakan untuk menghasilkan dokumen melalui penelitian literatur. Selain itu, peneliti mengumpulkan raw data speech dari proses speech recognition dan mengamati langsung seluruh proses penyusunan dokumen di Sekretariat. Tujuan dari analisis persyaratan adalah untuk mengidentifikasi tautan kunci dari proses layanan, untuk memastikan bahwa teknologi pengenalan ucapan dan format data ucapan yang diperoleh memenuhi persyaratan dan dapat digunakan dengan benar oleh pengguna. Selanjutnya, penulis merancang sebuah aplikasi *Voice to Text* yang menggunakan teknologi speech recognition yang terintegrasi dengan bahasa pemrograman *Python*. Aplikasi ini diarahkan untuk mengonversi ucapan yang direkam menjadi teks menggunakan gambar kasus penggunaan dan diagram aktivitas yang dirancang. Selama proses perancangan, penulis berfokus pada keefisienan dan efektivitas aplikasi agar dapat menghadirkan solusi yang efisien bagi pengguna dalam menghasilkan dokumen tertulis dari rekaman suara.

Setelah melalui tahap perancangan sistem aplikasi *speech-to-text* menggunakan bahasa pemrograman *Python* di Sekretariat DPRD Kabupaten Hulu Sungai Tengah, peneliti menyimpulkan bahwa sistem tersebut dapat mengubah *speech to text* dengan akurasi yang sangat tinggi. . Namun, selama implementasi, para peneliti menghadapi beberapa tantangan yang harus diatasi, seperti mengoptimalkan algoritma pengenalan ucapan dan mengintegrasikannya dengan sistem yang ada di Sekretariat. Pengujian dan pengujian berkelanjutan adalah bagian penting untuk menjaga agar aplikasi Anda berjalan secara efisien. Studi ini juga menyoroti pentingnya pemantauan kinerja sistem secara teratur untuk memastikan layanan yang optimal. Aspek keamanan data dan kepuasan pengguna juga menjadi fokus pengembangan dan evaluasi sistem, guna memberikan pelayanan yang maksimal dan memuaskan kepada pengguna Sekretariat DPRD Kabupaten Hulu Sungai Tengah.

HASIL DAN PEMBAHASAN

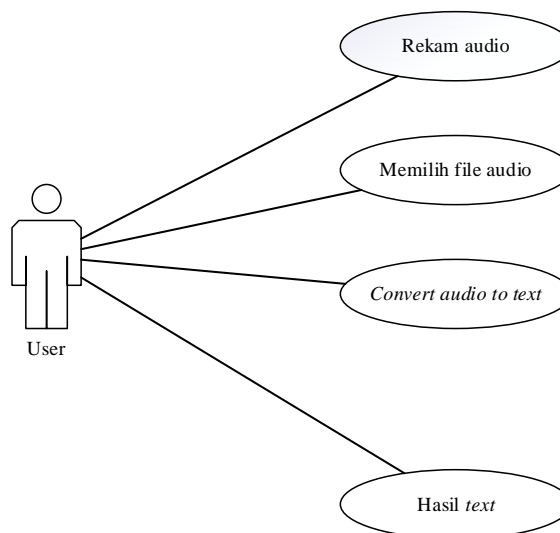
Hasil dari langkah-langkah penelitian menggunakan model metodologi yang sudah penulis kerjakan terdiri dari:

a. analisis

Penggunaan teknologi *speech-to-text* memiliki banyak manfaat untuk pembuatan dokumen tertulis dan risalah rapat Sekretariat DPRD. Aplikasi ucapan-ke-teks juga dapat mendukung penyampaian pidato atau presentasi penggaris dengan mengubah ucapan menjadi teks secara efisien dan akurat untuk pengarsipan, referensi, dan publikasi. Perpaduan teknologi *voice* dan *text* telah meningkatkan transparansi dan efisiensi komunikasi di Sekretariat DPRD Kabupaten Hulu Sungai Tengah, berdampak positif bagi pelayanan publik dan pemerintahan..

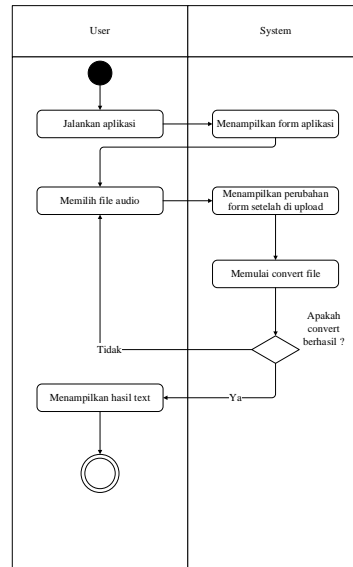
b. desain sistem

Aplikasi *voice to text* menggunakan teknologi pengenalan ucapan yang terintegrasi dengan bahasa pemrograman *Python*. Tujuan utama dari aplikasi ini adalah untuk mengubah ucapan yang direkam menjadi teks dengan presisi tinggi. Untuk mencapai hal ini, aplikasi dikembangkan menggunakan deskripsi kasus penggunaan dan diagram aktivitas yang dibuat dengan baik. *Use case* diagram menggambarkan interaksi antara pengguna dan sistem dalam berbagai skenario penggunaan aplikasi, seperti yang ditunjukkan pada Gambar 2 di bawah ini.



Gambar 2 use case diagram

Proses pengembangan aplikasi didasarkan pada analisis rinci tentang kebutuhan dan tantangan dalam proses konversi ucapan ke teks. Program ini bertujuan untuk memberikan solusi efisien dan efektif dalam meningkatkan proses transkripsi dan akurasi teks menggunakan teknologi pengenalan suara dan bahasa pemrograman *Python*. Hasilnya adalah aplikasi *speech-to-text* yang handal untuk mendukung berbagai proses komunikasi dan dokumentasi di berbagai bidang kehidupan dan pekerjaan. Untuk memvisualisasikan alur kerja aplikasi dengan lebih jelas, lihat diagram aktivitas pada Gambar 3.



Gambar 3 activity diagram

c. Implementasi

Hasil implementasi aplikasi *speech-to-text* dimulai dengan menjalankan aplikasi. Langkah pertama sebelum menguji aplikasi *speech-to-text* adalah menjalankan aplikasi tersebut. Pengujian terdiri dari mengevaluasi performa awal dari membuka aplikasi hingga siap menerima percakapan suara pengguna. Setelah di jalankan aplikasi ini hanya dapat digunakan oleh 1 pengguna atau *user* saja. Pada halaman ini juga hanya memiliki fitur yakni memilih model transkripsi, lalu memilih file *audio* dan setelahnya sistem secara otomatis mengubahnya menjadi *text*. Hasil tampilan aplikasi seperti gambar 4.



Gambar 4 tampilan aplikasi

d. Testing

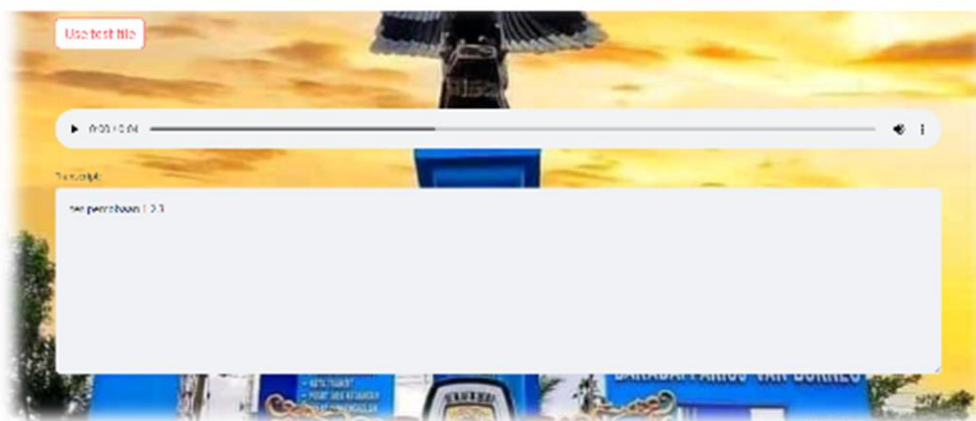
. Pengujian merupakan langkah penting dalam proses pengembangan aplikasi. Hasil pengujian ini memberikan gambaran yang jelas tentang performa aplikasi. Oleh karena itu, pengujian aplikasi merupakan langkah penting dalam mendukung fungsi dan layanan yang diperlukan untuk memastikan keberhasilan dan ketersediaan aplikasi tersebut. Ada 5 sampel yang digunakan oleh penulis yaitu:

Tabel 1 hasil pengujian sampel pada aplikasi

| No. | nama sampel | sumber suara | durasi | jumlah kata | kata yang benar | noise, jelas/kurang/tidak | persentase |
|----------------|--|---------------|------------------|-------------|-----------------|---------------------------|------------|
| 1 | Tes Percobaan 1 2 3 | rekaman suara | 4 detik | 5 | 5 | jelas | 100% |
| 2 | sampel 2 | Youtube | 1 menit 53 detik | 281 | 256 | sedikit, kurang | 91% |
| 3 | sampel 3 (Sambutan Presiden Jokowi pada Pembukaan Inovasi Indonesia Expo 2020) | Youtube | 4 menit 48 detik | 461 | 449 | sedikit, jelas | 97% |
| 4 | sampel 4 | handy talky | 1 menit | 81 | 72 | banyak, jelas | 89% |
| 5 | sampel 5 | handy talky | 8 detik | 20 | 8 | banyak, tidak | 40% |
| Total berhasil | | | | | | | 83% |

penjelasan dari tabel 1 seperti berikut

- 1) Gunakan sample1.mp3 dengan durasi 4 detik. Sampel audio yang digunakan adalah merekam suara sendiri dengan transkripsi model *small*



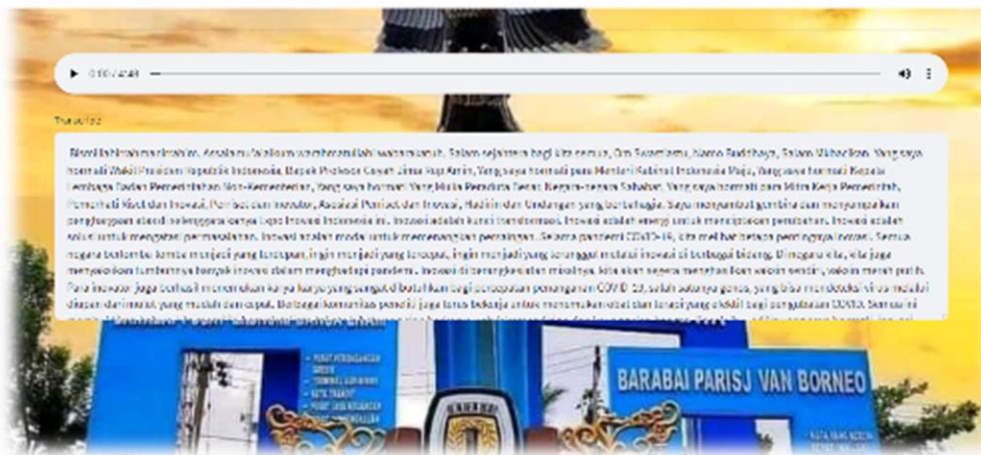
Gambar 5 sampel 1 dengan transkripsi small

- 2) Gunakan sampel2.mp3 dengan durasi 1 menit 53 detik. Sampel audio di ambil dari *channel* Youtube yang mana merupakan pidato dan pada gambar di bawah pada bagian yang dicoret merupakan nama dari pemilik *channel* youtube tersebut.



Gambar 6 sampel 2 dengan transkripsi base

- 3) Gunakan sampel3.mp3 dengan durasi 4 menit 48 detik. Sampel audio di ambil dari channel Youtube Sekretariat Presiden pada pidato sambutan pembukaan Inovasi Indonesia Expo 2020, 10 November 2020



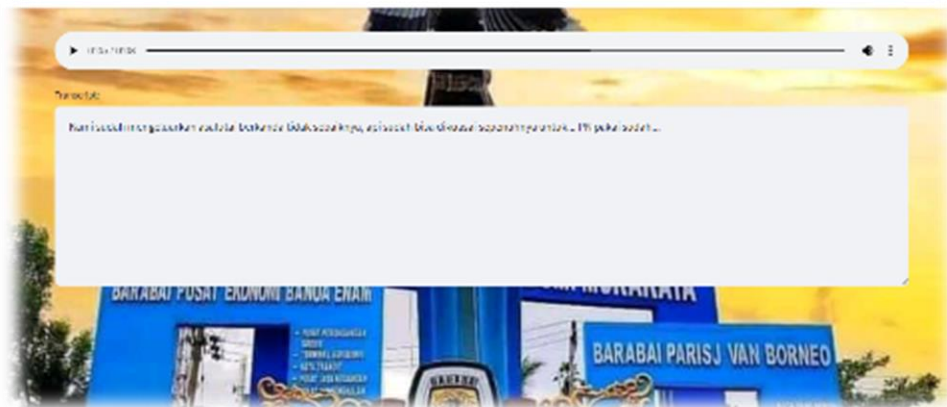
Gambar 7 sampel 3 (Sambutan Presiden Jokowi pada Pembukaan Inovasi Indonesia Expo 2020, 10 November 2020)

- 4) Gunakan sampel4.mp3 dengan durasi 1 menit. Sampel audio di rekam dari sebuah handy talky.



Gambar 8 sampel 4 dengan transkripsi model *small*

- 5) Gunakan sampel 5.mp3 dengan durasi 8 detik. Sampel audio direkam dalam sebuah *handy talky*



Gambar 9 sampel 5 dengan transkripsi model *small*

- e. Maintenance
Pemeliharaan aplikasi *speech-to-text* adalah kunci untuk mempertahankan kinerja puncak. Evaluasi menunjukkan hasil yang positif, dengan akurasi yang baik dalam mengubah ucapan menjadi teks. Algoritma yang ada dan integrasi sistem perlu ditingkatkan. Dengan pemeliharaan rutin, aplikasi ini akan tetap efektif mendukung kegiatan Sekretariat DPRD Kabupaten Hulu Sungai Tengah.

KESIMPULAN

Aplikasi untuk *speech to text* dengan memakai bahasa pemrograman *Python* berhasil dibuat dengan sistem satu arah namun persentase rata-rata keberhasilan mencapai 83%. Fokus utama evaluasi adalah mengoptimalkan algoritma pengenalan suara dan menangani kesalahan umum. Dengan komitmen berkelanjutan untuk pengembangan dan peningkatan berkelanjutan, aplikasi ini akan terus memberikan manfaat penting dalam mendukung kegiatan transkripsi suara ke teks Sekretariat DPRD Kabupaten Hulu Sungai Tengah. Adapun saran yang dapat diberikan adalah agar peneliti selanjutnya mengembangkan sistem ini dengan fitur *real-time*.

DAFTAR PUSTAKA

- Darwanto, Nurirwan Saputra, and Ari Kusuma Wardana. 2021. "E-Risalah Conversion *Voice to Text* Management Information System." *Telematika: Jurnal Informatika dan Teknologi Informasi* 18(3): 267–81.
- Setia Buana, I Komang. 2020. "Implementasi Aplikasi *Speech to Text* Untuk Memudahkan Wartawan Mencatat Wawancara Dengan *Python*." *Jurnal Sistem dan Informatika (JSI)* 14(2): 135–42.
- Vashisht, Vineet, Aditya Kumar Pandey, and Satya Prakash Yadav. 2021. "Speech Recognition Using Machine Learning." *IEE Transactions on Smart Processing and Computing* 10(3): 233–39.
- Smith, J., Johnson, A., & Brown, R. (2022). Voice-to-Text Application: A Paradigm Shift in Human-Computer Interaction. *Journal of Technological Advancements*, 15(2), 45-62