

1.2. TUJUAN DAN MANFAAT PENELITIAN

Tujuan dari penelitian ini adalah menganalisa beberapa *platform* AI *voice-over* dalam proses perancangan produksi konten informasi digital, serta menentukan platform yang tepat untuk digunakan dalam produksi konten informasi digital berbasis AI *voice over*.

2. KAJIAN LITERATUR

2.1. PENELITIAN TERDAHULU

Penelitian oleh Ren et al. (2020) mengenai FastSpeech 2: Fast and High-Quality End-to-End Text-to-Speech menjadi salah satu dasar penting dalam memahami perkembangan teknologi text-to-speech (TTS) modern. Studi ini menunjukkan bahwa model FastSpeech 2 mampu menghasilkan suara sintetis dengan kualitas tinggi dan proses inferensi yang sangat cepat karena bekerja secara non-autoregresif serta memiliki kontrol terhadap pitch, durasi, dan energi suara. Temuan tersebut menjelaskan bagaimana teknologi TTS dapat memangkas waktu produksi audio secara signifikan, sehingga mendukung kebutuhan produksi konten digital yang menuntut output cepat dan konsisten. Kelebihan utama penelitian ini terletak pada stabilitas dan fleksibilitas pengaturan prosodi, meskipun masih memiliki keterbatasan dalam ekspresi emosional suara. Relevansinya dengan penelitian ini adalah memberikan landasan teknis bahwa AI *voice-over* FastSpeech 2 dapat menghasilkan suara natural dan efisien.

Selanjutnya, penelitian yang dilakukan oleh Mubarak dan Santoso (2023) berjudul “Pemanfaatan Teknologi AI *Voice-Over* dalam Produksi Konten Digital Kreatif”. Penelitian tersebut mengkaji peran AI *Voice-Over* sebagai alternatif produksi suara yang lebih efisien dibandingkan metode konvensional, khususnya melalui pemanfaatan fitur-fitur kustomisasi yang disediakan platform berbasis AI, seperti pengaturan intonasi, emosi, tempo bicara, serta ketersediaan beragam karakter suara. Hasil penelitian menunjukkan bahwa kelengkapan fitur dan variasi suara memungkinkan satu sistem AI *Voice-Over* digunakan untuk berbagai jenis

konten, sehingga dapat mengurangi kebutuhan pengisi suara yang berbeda dan mempercepat proses produksi sekaligus menekan biaya. Relevansinya dengan adalah terletak pada kesamaan fokus terhadap penggunaan *AI Voice-Over* dalam produksi konten digital.

Perbedaan penelitian ini dengan penelitian terdahulu adalah penelitian terdahulu lebih berfokus pada aspek teknis pengembangan model TTS serta pemilihan *AI voice-over* dalam strategi konten digital. Sementara itu, penelitian ini berfokus pada menganalisa beberapa *platform AI voice-over* dalam proses perancangan produksi konten digital di Indonesia.

2.2. LANDASAN TEORI

Landasan teori merupakan pijakan konseptual yang memberikan arah analitis bagi penelitian ini. Pada penelitian mengenai integrasi teknologi Artificial Intelligence (AI) untuk pembuatan *voice-over* dalam produksi konten digital, pemahaman terhadap teori-teori dasar yang relevan menjadi penting. Teori tersebut tidak hanya menjelaskan fenomena yang dikaji, tetapi juga berfungsi sebagai perangkat analisis dalam menafsirkan temuan penelitian. Beberapa teori yang digunakan meliputi teori Text-to-Speech (TTS) modern, teori otomatisasi produksi konten digital, dan teori strategi konten digital. Ketiga kelompok teori ini saling melengkapi dan menopang keseluruhan argumentasi ilmiah penelitian

2.2.1 Platform AI Voice Over

ElevenLabs, Voiceover.com dan VEED.io merupakan platform TTS berbasis kecerdasan buatan yang saat ini banyak digunakan dalam industri kreatif, media digital, dan pendidikan. Teknologi ini mampu mengubah teks menjadi suara yang terdengar natural, seolah-olah diucapkan oleh narator profesional. Keunggulan utamanya terletak pada kualitas suara yang realistis, fleksibilitas dalam memilih karakter suara, serta kemampuan menghasilkan audio dengan latensi rendah. Hal ini menjadikan Platform *AI Voice Over* sebagai solusi praktis untuk kebutuhan *voice-over* dalam produksi video digital, audiobook, maupun konten edukasi.

Selain menghasilkan suara yang natural, Platform ini juga menawarkan fitur lanjutan seperti *voice cloning* dan kontrol emosional suara. Dengan *voice cloning*, pengguna dapat membuat suara sintetis yang menyerupai suara tertentu, sementara fitur kontrol emosional memungkinkan pengaturan intonasi, ekspresi, dan nuansa suara sesuai kebutuhan narasi. Fitur-fitur ini memperkuat relevansi Platform ini sebagai contoh nyata penerapan teori TTS modern, karena tidak hanya menghasilkan suara yang natural, tetapi juga memberikan fleksibilitas kreatif bagi pengguna.

Dalam konteks penelitian ini, Platform AI *voice over* menjadi representasi bagaimana teori TTS modern diterapkan secara praktis dalam industri kreatif. Jika penelitian akademis seperti Ren et al. (2020) menekankan aspek teknis model deep learning untuk menghasilkan suara stabil dan konsisten, maka AI *Voice Over* menunjukkan bagaimana teknologi tersebut diimplementasikan untuk mendukung produksi konten digital. Dengan demikian, Platform AI *Voice over* dapat diposisikan sebagai jembatan antara teori TTS modern dan kebutuhan nyata industri media, yaitu menghasilkan *voice-over* yang cepat, murah, dan berkualitas tinggi.

2.2.2 AI Voice Over

AI *Voice-Over* merupakan pengembangan dari teori *voice-over* konvensional yang mengintegrasikan kecerdasan buatan dalam proses produksi suara untuk kebutuhan konten audiovisual. Teknologi ini memanfaatkan sistem text-to-speech (TTS) yang mampu mengubah teks menjadi suara sintetis dengan karakteristik menyerupai suara manusia. Pemanfaatan AI dalam produksi *voice-over* memungkinkan suara dihasilkan secara lebih stabil dan seragam, sehingga dapat mendukung kebutuhan konten digital yang menuntut kualitas audio yang konsisten (Mubarak & Santoso, 2023).

Dalam produksi konten digital, teori *AI Voice-Over* menekankan prinsip efisiensi dan penyederhanaan alur kerja. Penggunaan *AI Voice-Over* mengurangi ketergantungan terhadap studio rekaman dan pengisi suara manusia, karena perubahan pada naskah dapat langsung dikonversi menjadi audio baru tanpa proses perekaman ulang. Kondisi ini menjadikan *AI Voice-Over* sebagai solusi praktis dalam produksi konten digital yang bersifat cepat, dinamis, dan membutuhkan fleksibilitas tinggi dalam proses revisi (Mubarak & Santoso, 2023).

Selain efisiensi produksi, teori *AI Voice-Over* juga menyoroti pentingnya konsistensi suara sebagai bagian dari strategi komunikasi dan pembentukan identitas merek (*brand voice*). Konsistensi karakter suara membantu audiens mengenali dan mengingat sebuah konten, terutama pada produksi konten berseri, media edukasi, dan konten pemasaran digital. Dengan kemampuan menghasilkan suara yang seragam dalam volume besar dan waktu singkat, *AI Voice-Over* berperan strategis dalam pengelolaan konten digital modern yang menekankan kesinambungan narasi dan kejelasan pesan (Mubarak & Santoso, 2023).

2.2.3 Konten Informasi Digital

Konten informasi digital merupakan media berbasis teknologi digital yang menyajikan pesan informatif melalui berbagai format, seperti *audiovisual*. Konten ini dirancang untuk memudahkan audiens dalam mengakses, memahami, dan menerima informasi secara cepat, luas, dan efisien seiring dengan perkembangan teknologi komunikasi digital.

Perkembangan kecerdasan buatan mendorong pemanfaatan *AI voice over* dalam produksi konten informasi digital, khususnya pada konten berbasis *audiovisual*. Chen & Harada (2023) menjelaskan bahwa teknologi *AI voice over* memungkinkan sistem menghasilkan suara sintetis yang mampu meniru karakter suara manusia secara natural, stabil, dan konsisten, sehingga mendukung penyampaian narasi yang lebih efektif.

Penggunaan AI *voice over* berpengaruh terhadap efektivitas penyampaian informasi karena kualitas suara menjadi faktor penting dalam membangun kejelasan pesan, daya tarik konten, serta persepsi audiens terhadap kredibilitas informasi. Dengan suara yang terdengar alami dan konsisten, konten digital berbasis teknologi mampu meningkatkan keterlibatan audiens dan memperkuat kepercayaan terhadap pesan yang disampaikan.

Tabel 2. 1. Perbedaan penelitian terdahulu dan landasan teori. Dokumentasi Pribadi

Penelitian terdahulu	Landasan teori
Penelitian Ren et al. (2020) mengenai <i>FastSpeech 2: Fast and High-Quality End-to-End Text-to-Speech</i> , yang membahas model TTS non-autoregresif untuk menghasilkan suara berkualitas tinggi dengan kontrol <i>pitch</i> , durasi, dan energi.	Platform AI <i>Voice-Over</i> : ElevenLabs, Voiceover.com, dan VEED.IO adalah platform TTS modern berbasis AI yang menghasilkan suara natural dengan fitur seperti voice cloning dan kontrol prosodi , sehingga relevan sebagai penerapan teori TTS dalam industri kreatif dan media digital
Artikel Mubarak & Santoso (2023) <i>Pemanfaatan Teknologi AI Voice-Over dalam Produksi Konten Digital Kreatif</i> yang mengkaji penggunaan AI <i>Voice-Over</i> sebagai alternatif produksi suara dalam industri kreatif digital	Penelitian Mubarak & Santoso (2023) meneliti pemanfaatan AI <i>Voice-Over</i> sebagai alternatif produksi suara dalam konten digital kreatif. Teori AI <i>Voice-Over</i> menekankan konsistensi suara, serta fleksibilitas penggunaan <i>voice-over</i> berbasis AI dalam berbagai bentuk konten digital.