

1. LATAR BELAKANG

Dunia informasi dan komunikasi saat ini dipenuhi oleh konten digital, terutama dalam bentuk video yang mendominasi berbagai platform seperti TikTok, Instagram, YouTube, dan layanan *streaming*. Perkembangan industri konten digital di Indonesia menunjukkan peningkatan yang signifikan, seiring dengan tingginya konsumsi video sebagai medium utama masyarakat digital (Faridatul & Purwanto, 2025). Kondisi ini mendorong kebutuhan produksi konten video yang cepat, variatif, dan konsisten, sehingga menuntut adanya proses kerja yang lebih efisien dalam industri kreatif dan teknologi. Salah satu elemen penting dalam produksi video adalah *voice-over*, yang berfungsi menyampaikan informasi, narasi, serta pesan teknologi secara jelas dan persuasif kepada audiens. Namun, produksi *voice-over* secara konvensional masih menghadapi kendala biaya *talent*, kebutuhan *studio*, serta proses revisi yang relatif panjang.

Tantangan tersebut semakin terasa ketika perusahaan dituntut untuk memproduksi konten digital dalam jumlah besar dan waktu yang singkat. Metode *voice-over* konvensional dinilai kurang fleksibel, terutama ketika terjadi perubahan naskah atau penyesuaian konsep *visual*, karena membutuhkan proses perekaman ulang yang memakan waktu dan biaya tambahan (Nirwana et al., 2023). Dalam konteks ini, pemanfaatan teknologi *Artificial Intelligence* (AI) *Voice-Over* menjadi solusi alternatif yang relevan. Teknologi ini memungkinkan teks diubah menjadi suara secara otomatis dengan kualitas yang semakin natural, didukung oleh perkembangan model *text-to-speech* (TTS) modern seperti *FastSpeech 2* yang mampu menghasilkan suara dengan prosodi dan stabilitas yang dapat dikontrol (Ren et al., 2020).

Penggunaan AI *Voice-Over* menjadi semakin relevan bagi perusahaan yang memproduksi konten informasi berbasis teknologi dalam format video tanpa melibatkan aktor atau *talent visual*. PT Teknologi Indonesia Group, sebagai perusahaan yang bergerak di bidang teknologi, membutuhkan konten informasi

digital yang menekankan kejelasan pesan teknis dan konsistensi narasi, tanpa ketergantungan pada kehadiran aktor di depan kamera. Oleh karena itu, *AI Voice-Over* dipilih sebagai solusi untuk menyampaikan informasi teknologi secara efektif melalui video berbasis *visual* dan animasi, sekaligus menjaga efisiensi produksi.

Berbagai platform *AI Voice-Over* seperti ElevenLabs, Voiceoover.com, dan VEED.IO menawarkan fitur yang semakin beragam, mulai dari pengaturan emosi, intonasi, kecepatan bicara, hingga teknologi *voice cloning*. Fitur-fitur tersebut memungkinkan perusahaan menyesuaikan karakter suara dengan kebutuhan konten dan identitas merek tanpa harus melakukan perekaman ulang secara manual. Selain itu, variasi suara yang tersedia memberikan fleksibilitas dalam produksi berbagai jenis konten teknologi, mulai dari video edukasi, penjelasan produk, hingga materi pemasaran digital (Mubarak & Santoso, 2023).

Dalam proses produksi, penulis mulai menguji dan menerapkan beberapa platform *AI Voice-Over* untuk mendukung produksi konten digital informasinya. Pengujian dilakukan dengan membandingkan beberapa platform guna menilai kualitas suara, kecepatan proses, kemudahan penggunaan, ketentuan lisensi, serta biaya yang dibutuhkan. Langkah ini dilakukan agar perusahaan dapat menentukan platform *AI Voice-Over* yang paling sesuai dengan kebutuhan produksi konten teknologi berbasis video tanpa aktor, sekaligus mendukung efisiensi dan konsistensi konten dalam jangka panjang (Martinez & O'Neill, 2022; Chen & Harada, 2023).

1.1. RUMUSAN DAN FOKUS

Bagaimana analisis komparatif *platform AI voice-over* dapat digunakan dalam perancangan produksi konten informasi digital? Penelitian ini berfokus pada Perbandingan ketiga platform *AI voice-over*, yaitu ElevenLabs, Voiceoover.com, dan VEED.IO, dalam mendukung perancangan produksi konten informasi digital.

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