

## BAB I

### PENDAHULUAN

#### 1.1 Latar Belakang

Desa Wisata Tigaraksa memiliki potensi besar sebagai destinasi wisata edukasi berbasis pertanian. Dengan lahan perkebunan yang luas, desa ini berupaya mengembangkan program edukasi lingkungan bagi anak-anak sekolah dasar untuk memperkenalkan flora, fauna, dan praktik pertanian berkelanjutan sesuai dengan penelitian [1]. Pemilihan tema flora dan fauna lokal sebagai fokus visual edukatif didasarkan pada dua pertimbangan utama: pertama, Desa Tigaraksa memiliki lahan perkebunan luas yang menjadi aset utama pengembangan desa wisata; kedua, tema flora dan fauna relevan untuk edukasi anak-anak terkait asal-usul makanan dan kearifan lokal. Dalam konteks *Community-Based Tourism (CBT)*, keterlibatan masyarakat lokal dalam menciptakan konten edukatif memiliki peran penting dalam membangun kemandirian desa serta memperkuat identitas lokal [2], [3].

Namun, implementasi program edukasi tersebut menghadapi kendala signifikan dalam penyediaan materi visual yang menarik dan kontekstual. Tantangan inilah yang mendorong eksplorasi solusi berbasis teknologi, di mana perkembangan *Artificial Intelligence (AI)*, khususnya *Generative AI* dengan model *text-to-image (T2I)*, menawarkan peluang inovatif. Teknologi T2I telah memasuki berbagai sektor kehidupan global, termasuk pendidikan, seni, dan pariwisata, karena kemampuannya menghasilkan visual dari prompt teks tanpa memerlukan keahlian desain grafis profesional [4], [5].

Berbagai penelitian menunjukkan bahwa penerapan T2I di bidang pendidikan dan desain dapat meningkatkan eksplorasi ide, memperkuat daya imajinasi, dan membantu pemahaman konseptual siswa melalui visualisasi interaktif [6]. Berdasarkan teori *Multimedia Learning*, pembelajaran menjadi lebih efektif ketika teks dan gambar disajikan secara bersamaan karena dapat memperkuat pemrosesan kognitif dan meningkatkan retensi informasi. Studi empiris menunjukkan bahwa penggunaan AI generatif dalam pembelajaran visual dapat meningkatkan keterlibatan siswa hingga 40% dan mempercepat proses iterasi desain hingga 3x lipat [7]. Teknologi ini mengubah cara manusia berkreasi dan belajar karena memungkinkan siapa pun—termasuk masyarakat lokal di desa wisata untuk menciptakan konten edukatif berkualitas secara mandiri.

Perkembangan teknologi *Text-to-Image Generative AI* menawarkan solusi potensial terhadap kendala tersebut. Melalui model seperti Gemini Nano Banana dan Stable Diffusion, masyarakat dapat menghasilkan gambar edukatif berkualitas tinggi hanya dengan menuliskan deskripsi teks yang relevan [8], [9]. Teknologi ini dapat membantu desa wisata menghasilkan konten pembelajaran secara mandiri serta memperluas akses pembelajaran berbasis visual yang kontekstual.



Meski demikian, implementasi T2I dalam konteks pendidikan menghadapi beberapa tantangan. Kualitas hasil visual sangat bergantung pada penyusunan prompt, dan hingga kini belum terdapat pedoman sistematis mengenai metode prompt yang paling efektif dalam menghasilkan visual edukatif yang sesuai dengan karakteristik pembelajaran anak-anak [10]. Berdasarkan kerangka *prompt engineering* yang dikembangkan, efektivitas prompt dipengaruhi oleh tiga dimensi utama yaitu spesifik deskripsi, struktur instruksi dan konteks komposisi visual. Namun, *framework* ini belum diaplikasikan secara sistematis dalam konteks pendidikan berbasis komunitas. Penelitian terdahulu masih banyak berfokus pada aspek teknis dan estetika, bukan pada konteks pembelajaran berbasis komunitas [11], [12]. Selain itu, perbandingan antar model T2I yang mempertimbangkan konteks edukatif, gaya visual, dan aksesibilitas juga masih jarang dilakukan. Hal ini menyebabkan pengguna non-teknis seperti pendidik atau pengelola desa wisata kesulitan memilih model yang paling sesuai dengan kebutuhan dan sumber daya yang tersedia [13]. Kesenjangan ini menunjukkan perlunya penelitian yang menggabungkan pendekatan teknis dan kontekstual.

Untuk menempatkan *Generative AI* dalam konteks pembelajaran, penelitian ini mengacu pada dua kerangka konseptual yaitu *TPACK (Technological Pedagogical Content Knowledge)* dan *SAMR (Substitution, Augmentation, Modification, Redefinition)*. Dalam kerangka TPACK, model T2I berperan sebagai komponen *Technological Knowledge* yang mendukung *Content Knowledge* dan *Educational Knowledge* yaitu strategi belajar berbasis visual interaktif). Sedangkan melalui model SAMR, penggunaan T2I dalam penelitian ini telah mencapai tahap *Modification–Redefinition*, di mana teknologi tidak sekadar menggantikan alat tradisional, tetapi menciptakan pengalaman belajar baru yang lebih eksploratif dan kontekstual [14].

Penelitian ini dirancang untuk menjawab kesenjangan tersebut melalui pendekatan eksperimental-komparatif yang membandingkan tiga variasi metode *prompt engineering* yaitu *descriptive*, *instruction-based*, dan *compositional*. pada dua model T2I yaitu *Gemini Nano Banana* dan *Stable Diffusion*. Hasil visual yang dihasilkan dievaluasi secara kuantitatif menggunakan CLIPScore dan secara kualitatif melalui penilaian relevansi oleh pendidik dan masyarakat Desa Wisata Tigaraksa.

Penelitian ini memberikan kontribusi ganda. Secara teoretis, memperkaya literatur tentang *prompt engineering* untuk konteks pendidikan dan pariwisata berbasis komunitas [15]. Secara praktis, penelitian ini menghasilkan panduan implementasi Generative AI yang dapat digunakan oleh pengelola Desa Wisata Tigaraksa untuk memproduksi konten edukatif secara mandiri, efisien dan berkelanjutan. Dengan demikian, Generative AI tidak hanya menjadi inovasi teknologi, tetapi juga pengembangan wisata edukasi berbasis identitas lokal [16].

Berdasarkan penjelasan di atas, penelitian ini secara khusus berfokus pada evaluasi metode *prompt engineering* (*descriptive*, *instruction-based*, dan *compositional*) dalam menghasilkan visual edukatif untuk pembelajaran anak-anak sekolah dasar di Desa Wisata Tigaraksa.

Fokus komparatif penelitian ini terletak pada perbandingan dua model *Text-to-Image* dengan karakteristik berbeda: *Stable Diffusion* yang bersifat open-source dengan kontrol parameter tinggi dan *Gemini Nano Banana* yang berbasis multimodal transformer dengan kemampuan pemahaman bahasa indonesia dan juga natural. Dengan membandingkan kedua model tersebut menggunakan tiga metode prompt, penelitian ini tidak hanya mengukur performa teknis model AI, tetapi juga mengevaluasi relevansi dan kesesuaian hasil visual terhadap konteks edukatif lokal melalui kombinasi metrik kuantitatif (CLIPScore) dan penilaian kualitatif dari guru, pengelola desa dan anak-anak sebagai pengguna akhir.

Gap penelitian yang diisi oleh studi ini mencakup tiga aspek yang belum terjawab oleh penelitian terdahulu. Pertama, belum ada studi yang secara sistematis menghubungkan variasi metode prompt (*descriptive, instruction-based, compositional*) dengan kualitas hasil visual dalam konteks pendidikan komunitas pedesaan Indonesia. Kedua, mayoritas penelitian sebelumnya berfokus pada evaluasi teknis model T2I tanpa mempertimbangkan faktor aksesibilitas, dukungan bahasa lokal dan kelayakan adopsi oleh pengguna non-teknis seperti guru desa atau pengelola wisata. Ketiga, belum ditemukan framework evaluasi yang mengintegrasikan penilaian objektif (CLIPScore) dengan evaluasi manusia yang mempertimbangkan konteks sosial, budaya, dan edukatif lokal.

Dengan demikian, kontribusi penelitian ini bersifat ganda: secara teoretis memperkuat pengembangan kerangka prompt engineering dalam pendidikan , dan secara praktis menghasilkan panduan implementasi *Generative AI* yang dapat digunakan oleh Desa Wisata Tigaraksa untuk memproduksi konten edukatif secara efisien dan berkelanjutan.

Untuk mencapai tujuan tersebut, penelitian ini menggunakan pendekatan eksperimental komparatif yang terdiri dari tiga tahap utama yaitu pengembangan variasi prompt berdasarkan *framework prompt engineering*, generasi visual menggunakan dua model T2I dengan karakteristik berbeda dan evaluasi komprehensif melalui metrik kuantitatif dan penilaian manusia. Metodologi ini memungkinkan identifikasi strategi prompt yang paling efektif serta rekomendasi model yang paling sesuai untuk konteks Desa Wisata Tigaraksa.

## 1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, penelitian ini merumuskan dua pertanyaan penelitian sebagai berikut:

1. Bagaimana menentukan metode prompt (deskriptif, instruktif dan komposisi) memengaruhi kualitas hasil visual T2I dalam konteks edukasi anak dan promosi Desa Wisata Tigaraksa?

2. Rekomendasi model *text-to-image* *Gemini Nano Banana* atau *Stable Diffusion* untuk menghasilkan visual edukatif anak-anak di Desa Wisata Tigaraksa?

### **1.3 Batasan Masalah**

Untuk menjaga fokus dan ruang lingkup penelitian agar tetap terarah, maka batasan-batasan dalam penelitian ini adalah sebagai berikut:

1. Penelitian ini dibatasi pada pemanfaatan teknologi Text-to-Image (T2I) sebagai media pendukung kegiatan edukasi anak-anak sekolah dasar di Desa Wisata Tigaraksa, tanpa mencakup pengembangan model baru, pelatihan ulang (*retraining*), maupun modifikasi arsitektur algoritma.
2. Model T2I yang digunakan dalam penelitian ini dibatasi pada Gemini Nano Banana dan Stable Diffusion, dengan pertimbangan aksesibilitas, keterbukaan teknologi, serta relevansi terhadap konteks penelitian, di mana seluruh prompt disusun menggunakan Bahasa Indonesia untuk menguji kemampuan model dalam memahami konteks lokal dan bahasa non-Inggris.
3. Target audiens penelitian dibatasi pada anak-anak usia 7–12 tahun (kelas 1–6 Sekolah Dasar), yang berada pada tahap perkembangan kognitif operasional konkret, sehingga pembelajaran lebih efektif melalui representasi visual yang nyata dan kontekstual.
4. Variasi metode prompt yang diuji dibatasi pada tiga kategori, yaitu *Descriptive Prompt*, *Instruction Prompt*, dan *Compositional Prompt*, dengan tema fauna dan flora perkebunan.
5. Evaluasi hasil visual dibatasi pada penggunaan CLIPScore dan penilaian manusia (*human evaluation*) berbasis rubrik yang mencakup relevansi konteks lokal, akurasi representasi objek, kualitas estetika, serta kesesuaian dengan audiens anak-anak dengan melibatkan pendidik, pengelola desa wisata, dan praktisi.

### **1.4 Tujuan dan Manfaat Penelitian**

#### **1.4.1 Tujuan Penelitian**

Berdasarkan rumusan masalah di atas, penelitian ini bertujuan untuk:

1. Mengidentifikasi dan menganalisis pengaruh variasi metode prompt (deskriptif, instruktif dan komposisi) terhadap kualitas hasil visual yang dihasilkan oleh teknologi T2I, dengan fokus pada aspek relevansi konteks lokal dan kesesuaian untuk edukasi anak serta promosi Desa Wisata Tigaraksa.
2. Membandingkan dan mengevaluasi kinerja dua model T2I yaitu *Gemini* dan *Stable Diffusion* berdasarkan kualitas output visual yang dihasilkan untuk kebutuhan edukasi anak, serta merekomendasikan model yang paling optimal untuk implementasi praktis dengan mempertimbangkan aspek aksesibilitas, biaya dan kesesuaian output.

#### **1.4.2 Manfaat Penelitian**

1. Memperkaya literatur akademik tentang *prompt engineering* dalam teknologi *Text-to-Image* dengan fokus pada konteks pendidikan berbasis komunitas dan wisata edukasi, area yang masih terbatas dalam literatur yang ada.
2. Menyediakan panduan praktis yang komprehensif dan berbasis bukti bagi pengelola Desa Wisata Tigaraksa dalam memanfaatkan teknologi T2I untuk menghasilkan materi edukasi anak secara mandiri, efisien, dan berkelanjutan.
3. Menghasilkan prototipe visual dan template prompt yang dapat langsung digunakan atau diadaptasi oleh Desa Wisata Tigaraksa untuk berbagai keperluan edukasi.

### **1.5 Sistematika Penulisan**

#### **BAB 1 PENDAHULUAN**

Bab ini menguraikan latar belakang penelitian yang dimulai dari kondisi dan tantangan yang dihadapi Desa Wisata Tigaraksa dalam menyediakan materi visual edukatif dan promosi, potensi teknologi *Text-to-Image (T2I)* sebagai solusi alternatif, serta kesenjangan penelitian yang perlu dijembatani. Bab ini juga memuat rumusan masalah yang menjadi fokus penelitian, batasan masalah yang mendefinisikan ruang lingkup kajian secara rinci, tujuan

penelitian yang ingin dicapai, manfaat penelitian, serta sistematika penulisan yang menggambarkan struktur keseluruhan penelitian.

## BAB II LANDASAN TEORI

Bab ini menyajikan kajian literatur dan teori-teori yang menjadi landasan penelitian, meliputi konsep *Generative Artificial Intelligence (GenAI)* dan perkembangannya, teknologi *Text-to-Image (T2I)* dan mekanisme kerjanya, konsep dan praktik prompt engineering dengan berbagai pendekatannya, karakteristik dan perbandingan model-model T2I utama (*Gemini & Stable Diffusion*), penerapan T2I dalam konteks pendidikan dan pembelajaran visual, pemanfaatan T2I untuk promosi wisata dan konten kreatif serta penelitian-penelitian terdahulu yang relevan dengan topik ini. Landasan teori ini berfungsi sebagai kerangka konseptual untuk menganalisis dan menginterpretasi temuan penelitian.

## BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan pendekatan dan metode penelitian yang digunakan secara sistematis. Penelitian ini dimulai dengan kajian literatur untuk mengidentifikasi metode prompt yang efektif dalam konteks edukasi berdasarkan penelitian terdahulu. Selanjutnya, dipilih dua model T2I yaitu Gemini dan Stable Diffusion untuk dibandingkan kinerjanya dalam menghasilkan visual sesuai kebutuhan Desa Wisata Tigaraksa. Metrik evaluasi dirancang berdasarkan kerangka penilaian kualitas prompt dan visual dari literatur terdahulu. Bab ini menjelaskan secara rinci desain penelitian, tahapan pengembangan variasi prompt (deskriptif, naratif, dan instruktif), prosedur generasi visual, evaluasi yang terstruktur, serta teknik analisis untuk membandingkan kinerja kedua model. Dengan pendekatan ini, penelitian dapat direplikasi dan hasilnya dapat divalidasi secara transparan.

## BAB IV HASIL DAN PEMBAHASAN

Bab ini memaparkan hasil penelitian yang diperoleh dari eksperimen variasi prompt dan perbandingan model T2I secara rinci dan terstruktur. Bagian hasil

mencakup visualisasi dan dokumentasi hasil generasi gambar dari berbagai kombinasi prompt dan model, analisis deskriptif kualitas visual berdasarkan penilaian panel, identifikasi pola dan temuan utama terkait pengaruh metode prompt terhadap kualitas output, perbandingan kinerja ketiga model T2I berdasarkan berbagai aspek evaluasi. Bagian pembahasan menginterpretasikan temuan tersebut dalam konteks teori dan penelitian terdahulu, menjelaskan implikasi praktis untuk implementasi di Desa Wisata Tigaraksa, mengidentifikasi faktor-faktor yang memengaruhi keberhasilan dan kegagalan, serta merumuskan rekomendasi praktis berdasarkan bukti yang ditemukan.

## BAB V SIMPULAN DAN SARAN

Bab ini berisi simpulan yang menjawab rumusan masalah berdasarkan hasil penelitian secara ringkas dan jelas, meliputi kesimpulan tentang pengaruh variasi prompt terhadap kualitas visual, perbandingan kinerja kedua model T2I, dan rekomendasi model serta metode prompt terbaik untuk konteks Desa Wisata Tigaraksa. Bab ini juga memuat saran-saran yang terbagi menjadi beberapa kategori: saran untuk implementasi praktis di Desa Wisata Tigaraksa (langkah-langkah adopsi, pelatihan, dan keberlanjutan), saran untuk penelitian lanjutan (eksplorasi aspek yang belum terjangkau dalam penelitian ini), saran metodologis (perbaikan desain penelitian untuk studi serupa), serta rekomendasi kebijakan terkait pemanfaatan teknologi T2I dalam pendidikan. Keterbatasan penelitian juga dipaparkan secara transparan untuk memberikan konteks interpretasi hasil dan arah pengembangan penelitian selanjutnya.

UNIVERSITAS  
MULTIMEDIA  
NUSANTARA