

Teknik Prompt ANVS Progresif untuk Bing Image Creator dan Interpretasi Kesenjangan Gambar AI

Andreas James Darmawan^{1*}, Peter Min Hyuk Yoo², Rivandy Aprilian Hartawan³

Desain Komunikasi Visual, Jakarta International University

Email : ^{1*}james.dar@jiu.ac, ²minyoo@jiu.ac, ³rivandy23@jiu.ac

ABSTRAK: Penelitian ini bertujuan untuk membantu pengguna platform *AI image generator*, khususnya Bing Image Creator, dalam merancang *prompt* yang efektif untuk menghasilkan gambar sesuai kebutuhan. Metode penelitian bersifat kualitatif, dilakukan dengan mengembangkan dan menguji formula ANVS (*Adjective, Noun, Verb, Style*), yang dirancang untuk memandu pengguna dalam menyusun *prompt*. Formula ini diterapkan secara bertahap, dari yang sederhana hingga kompleks, untuk mengevaluasi bagaimana setiap komponen memengaruhi hasil gambar yang dihasilkan. Penelitian ini berargumen bahwa struktur *prompt* yang terorganisir, seperti yang diusulkan oleh ANVS, dapat meningkatkan akurasi dan relevansi gambar yang dihasilkan oleh *AI image generator*. Hasilnya menunjukkan bahwa penerapan formula ANVS mampu memberikan panduan praktis dalam menyusun *prompt*, dengan bukti berupa pencatatan hasil eksperimen menggunakan Bing *Image Creator*. Kesimpulannya, formula ANVS efektif dalam membantu pengguna menciptakan *prompt* yang lebih terarah dan efisien. Penelitian ini memberikan kontribusi signifikan bagi pengembangan teknik *prompt engineering* di era digital, terutama dalam memaksimalkan potensi *AI image generator*; Bing *Image Creator*.

Diterima:
02/09/2024
Direvisi:
12/10/2024
Disetujui:
03/12/2024

Kata kunci: AI Image Generator, Bing Image Creator, Prompt Engineering.

Progressive ANVS Prompts Engineering for Bing Image Creator and AI Image Gap Interpretation

ABSTRACT : *This study aims to help users of the AI image generator platform, especially Bing Image Creator, in designing effective prompts to produce images according to their needs. The research method is qualitative, conducted by developing and testing the ANVS (Adjective, Noun, Verb, Style) formula, which is designed to guide users in composing prompts. This formula is applied in stages, from simple to complex, to evaluate how each component affects the resulting image. This study argues that an organized prompt structure, as proposed by ANVS, can improve the accuracy and relevance of images generated by AI image generators. The results show that the application of the ANVS formula is able to provide practical guidance in composing prompts, with evidence in the form of recording experimental results using Bing Image Creator. In conclusion, the ANVS formula is effective in helping users create more focused and efficient prompts. This study makes a significant contribution to the development of prompt engineering techniques in the digital era, especially in maximizing the potential of AI image generators; Bing Image Creator.*

Keywords: AI Image Generator, Bing Image Creator, Prompt Engineering.

PENDAHULUAN

Dalam era digital yang semakin berkembang, penguasaan teknologi kecerdasan buatan (AI) menjadi salah satu kompetensi yang sangat penting, terutama di bidang kreatif dan desain. *AI Image Generator* adalah alat yang efektif dan populer, baik bagi desainer, seniman, maupun bagi para pengguna umum untuk menciptakan gambar yang sesuai dengan kebutuhan pengguna (Chiou, 2023). Teknologi ini tidak hanya mempercepat proses kreatif tetapi juga memungkinkan penciptaan visual yang tidak terbatas oleh keterampilan teknis tradisional (Beyan, 2023). Oleh karena itu, diperlukan pemahaman dan kemampuan *prompt engineering* untuk mengoperasikan *AI image generator* secara efektif menjadi semakin mendesak di masa sekarang.

Untuk mengerti teknik *prompt*, sangat penting bagi pengguna untuk memahami bagaimana cara *AI image generator* mengerti perintah dalam bentuk teks (*prompt*), dan dalam penelitian ini, menggunakan AI bernama *Bing Image Creator*, menangkap dan menginterpretasikan *prompt* yang pengguna buat (Liu, 2022).

Memahami mekanisme internal AI dalam menterjemahkan deskripsi tekstual menjadi visual adalah kunci untuk menghasilkan gambar yang lebih akurat dan memuaskan (Hermanto, 2024). Ini memerlukan pendekatan yang lebih strategis dalam pembuatan *prompt engineering*, di mana setiap elemen deskripsi diperhatikan secara detail untuk memaksimalkan kualitas hasil akhir.

Teknik *prompt* adalah pendekatan terstruktur dalam menyusun instruksi teks untuk memandu AI image generator menghasilkan gambar sesuai kebutuhan pengguna. perancangan dan penyesuaian prompt dapat meningkatkan kualitas gambar yang dihasilkan oleh model generatif seperti *Stable Diffusion*, termasuk pendekatan reinforcement learning untuk meningkatkan hasil estetis dan keselarasan dengan niat awal pengguna (Hao, 2023). Teknik ini bertujuan untuk memaksimalkan potensi teknologi dengan memanfaatkan kombinasi kata dan frasa yang dirancang secara strategis, memastikan hasil gambar yang lebih akurat dan relevan.

Formula ANVS, yaitu: *Adjective* (kata keterangan) + *Noun* (kata benda) + *Verb* (kata kerja) + *Style* (gaya visual) adalah formula yang dirancang untuk mengoptimalkan pembuatan *prompt* secara bertahap. Pendekatan berbasis fine-tuning yang memanfaatkan *reinforcement learning* (penguatan dalam pembelajaran) untuk menghasilkan prompt yang lebih efektif dan adaptif, relevan dengan prinsip pengembangan ANVS progresif (Schuhmann, 2023). Dengan memulai dari struktur sederhana hingga yang lebih kompleks, ANVS membantu pengguna mengeksplorasi pengaruh setiap elemen terhadap hasil gambar yang dihasilkan. Pendekatan ini memberikan panduan sistematis untuk mengembangkan keterampilan prompt engineering.

Bing *Image Creator* sendiri adalah platform berbasis kecerdasan buatan yang memungkinkan pengguna menciptakan gambar berdasarkan prompt teks. Alat ini memanfaatkan teknologi AI untuk menginterpretasikan instruksi pengguna dan menghasilkan visual yang mendekati deskripsi yang diberikan. Eksplorasi penggunaan Bing *Image Creator* dalam menghasilkan gambar berbasis teks, memerlukan adaptasi prompt dalam memperbaiki pengalaman pengguna. (Wang,

2023). Platform ini menjadi fokus penelitian karena popularitasnya di kalangan desainer dan pengguna umum.

Penelitian ini menunjukkan bahwa struktur prompt yang terarah, seperti formula ANVS, mampu meningkatkan kualitas hasil gambar. Hasil desain dari interpretasi prompt dapat memengaruhi hasil visual yang dihasilkan, memberikan wawasan tentang pendekatan evaluasi untuk penelitian berbasis *prompt* (Oppenlaender, 2023). Dengan mendokumentasikan proses dan hasil eksperimen, penelitian ini menyoroti pentingnya pendekatan sistematis dalam *prompt engineering*, memberikan kontribusi bagi pengembangan teknik baru yang relevan di era digital.

Kesenjangan gambar AI sendiri merujuk pada perbedaan antara hasil gambar yang dihasilkan oleh AI image generator dengan harapan pengguna. Kesenjangan estetis dan semantis dalam gambar AI memerlukan strategi adaptasi pembelajaran yang berbasis pengalaman, untuk menjembatani kesenjangan dalam menghasilkan desain dari *prompt* yang lebih presisi (Xu, 2023). Fenomena ini terjadi karena kurangnya pemahaman pengguna tentang cara menyusun prompt yang efektif atau keterbatasan algoritma AI. Penelitian ini mencoba menjembatani kesenjangan tersebut melalui pengembangan formula ANVS.

Meskipun demikian, masih banyak pengguna yang menghadapi tantangan dalam mendapatkan hasil gambar yang sesuai dengan harapan pengguna (Noel, 2024). Banyak pengguna merasa frustrasi karena hasil gambar yang dihasilkan oleh AI *image generator* sering kali tidak sesuai dengan deskripsi yang diberikan dalam *prompt engineering* (McGee, 2023). Hal ini disebabkan oleh kurangnya pemahaman tentang *prompt engineering* yaitu, bagaimana AI *image generator* menganalisis dan memproses *input* yang diberikan. Ketidakpastian ini menimbulkan kesenjangan antara ekspektasi dan realitas, yang pada akhirnya dapat mengurangi efektivitas penggunaan teknologi ini. Semua kemajuan teknologi sebaiknya dipelajari sejak dini (Darmawan, 2024).

Dalam penelitian ini, penulis mengembangkan dan menguji sebuah formula bernama ANVS, yaitu: *Adjective* (kata keterangan) + *Noun* (kata benda) + *Verb*

(kata kerja) + *Style* (gaya visual), yang dirancang untuk membantu pengguna dalam merancang *prompt engineering* yang tepat sehingga menghasilkan gambar yang sesuai pengguna (Allingham, 2023). Dengan menggunakan formula *prompt engineering* ini, penulis secara bertahap mengembangkan *prompt*, dimulai dari yang sederhana hingga cukup kompleks, untuk mengeksplorasi bagaimana setiap komponen dalam ANVS dalam mempengaruhi *output* hasil gambar yang dihasilkan oleh Bing *Image Creator*. Tingkat kompleksitas *prompt* didasari oleh tingkat detail hasil gambar yang hendak dicapai, hal ini membutuhkan kemampuan literasi yang deskriptif dalam menerjemahkan *image-to-texts* pada *prompt* (Crowson, 2022). Temuan dari penelitian ini diharapkan dapat memberikan panduan praktis bagi pengguna untuk menghasilkan gambar yang lebih sesuai dengan harapan pengguna.

METODE

Penelitian ini menggunakan metode kualitatif deskriptif yang bertujuan untuk menggambarkan dan menganalisis proses pengembangan *prompt* dalam menghasilkan gambar yang sesuai dengan harapan menggunakan Bing Image Creator. Pendekatan ini dipilih untuk memahami secara mendalam bagaimana AI *image generator* ini bekerja dan bagaimana formula ANVS (*Adjective + Noun + Verb + Style*) dapat membantu dalam optimalisasi pembuatan *prompt engineering*.

Tabel 1. Langkah Penelitian.
Sumber: Darmawan, 2024.

Aspek	Deskripsi
Tahap 1: Pemilihan Platform	<ul style="list-style-type: none"> • Dimulai dengan memilih platform Bing Image Creator. • Perbandingan dilakukan terhadap Getty Generative AI, Stable Diffusion, Chat GPT DALL-E, Adobe Firefly, dan Midjourney untuk menilai kelebihan dan kekurangan masing-masing.
Tahap 2: Pengembangan <i>Prompt Engineering</i> Berbasis Elemen ANV (<i>Adjective + Noun + Verb + Style</i>)	<ul style="list-style-type: none"> • Pembuatan dan pengujian elemen ANV (<i>Adjective + Noun + Verb</i>). • Menggunakan <i>prompt</i> sederhana dan dievaluasi secara progresif.
Tahap 3: Penambahan Elemen <i>S (Style)</i> pada <i>Prompt Engineering</i> .	<ul style="list-style-type: none"> • Menambahkan elemen S (<i>Style</i>) dalam <i>prompt</i> untuk mengeksplorasi pengaruh gaya visual. • Evaluasi dilakukan untuk setiap perubahan dalam hasil gambar.

Tahap 4: Analisis dan Kesimpulan

- Analisis berdasarkan proses pengembangan prompt.
 - Mengidentifikasi bagaimana Bing Image Creator menangkap deskripsi visual.
 - Mengkaji regulasi atau kendala yang membatasi hasil gambar.
-

Pengumpulan data dilakukan dengan metode snowball, penelitian dimulai dengan pemilihan platform Bing Image Creator setelah melakukan perbandingan dengan beberapa AI *image generator* lainnya, yaitu Getty Generative AI, Stable Diffusion, Chat GPT DALL-E, Adobe Firefly, dan Midjourney. Perbandingan ini dilakukan untuk menilai kelebihan dan kekurangan masing-masing *platform*, yang kemudian mengarahkan pada pemilihan Bing Image Creator sebagai fokus penelitian karena pertimbangan kemudahan akses, kualitas gambar, dan relevansi penggunaannya dalam studi ini.

Selanjutnya, pembuatan *prompt* dilakukan dengan menggunakan formula ANV (*Adjective + Noun + Verb*). Pada tahap ini, setiap *prompt engineering* sederhana diuji coba dan hasil gambar yang dihasilkan dievaluasi secara progresif. Setiap tahap pengujian bertujuan untuk mengidentifikasi bagaimana setiap elemen deskriptif ANV mempengaruhi hasil visual yang dihasilkan oleh AI *image generator* (McGee, 2023). Setelah beberapa iterasi, kemudian ditambahkan elemen S (*Style*) dalam *prompt engineering* untuk mengeksplorasi bagaimana gaya visual tertentu dapat dikontrol dan dihasilkan oleh Bing Image Creator.

Penelitian ini berakhir dengan menarik kesimpulan berdasarkan proses pengembangan prompt yang dilakukan. proses analisis *prompt engineering* dari Bing Image Creator dalam menangkap dan menerjemahkan deskripsi visual yang diberikan pengguna, serta mengidentifikasi beberapa kendala atau regulasi *prompt engineering* yang membatasi kemampuan AI dalam menghasilkan gambar tertentu (Al Saud, 2024). Hasil penelitian ini diharapkan dapat memberikan panduan praktis bagi pengguna dalam membuat *prompt engineering* yang lebih efektif dan memahami keterbatasan AI *image generator*, khususnya Bing Image Creator.

HASIL DAN PEMBAHASAN

1. Pemilihan Platform

Pengumpulan data dalam penelitian ini dilakukan dengan metode snowball sampling, yang dimulai dengan pemilihan platform Bing Image Creator sebagai objek studi utama. kemudian dilakukan perbandingan terhadap beberapa AI image generator lainnya seperti Getty Generative AI, Stable Diffusion, Chat GPT DALL-E, Adobe Firefly, dan Midjourney untuk mengidentifikasi kelebihan dan kekurangan masing-masing. Melalui perbandingan ini terkumpul data dan wawasan, tentang bagaimana masing-masing platform bekerja dalam menghasilkan gambar berdasarkan *prompt* yang diberikan (McGee, 2023). Dengan terus memperluas lingkup penelitian ke platform lain yang relevan, penguasaan pembuatan *prompt* dapat terbangun lebih menyeluruh dan mendalam tentang kekuatan, kelemahan, dan karakteristik unik dari setiap AI *image generator* yang dievaluasi.

Tabel 2. Kelebihan dan Kekurangan AI Image Generation.
Sumber: Darmawan, 2024.

AI Image Generator	Kelebihan	Kekurangan
Getty Generative AI	<ul style="list-style-type: none"> • Kualitas gambar sangat tinggi dengan akses ke koleksi gambar Getty Images yang luas. • Legalitas dan hak cipta lebih aman karena berbasis dan sinkron dengan koleksi gambar Getty Images. • Fokus pada kebutuhan industri kreatif. 	<ul style="list-style-type: none"> • Terbatas pada pengguna profesional dan berlisensi. • Biaya yang tinggi untuk akses penuh. • Pilihan kreatif terbatas pada konten Getty Images.
Bing Image Creator	<ul style="list-style-type: none"> • Gratis dan mudah diakses melalui Bing. • Integrasi dengan pencarian web Bing memudahkan inspirasi. • <i>User-friendly</i> untuk pemula. 	<ul style="list-style-type: none"> • Kualitas gambar bisa lebih rendah dibandingkan generator lainnya. • Pilihan dan kontrol kustomisasi terbatas. • Hasil terlalu generik.
Stable Diffusion	<ul style="list-style-type: none"> • <i>Open-source</i> dan dapat di-hosting secara lokal, memberi kontrol penuh kepada pengguna. 	<ul style="list-style-type: none"> • Memerlukan setup teknis yang kompleks dan pengetahuan pemrograman.

	<ul style="list-style-type: none"> • Sangat <i>customizable</i> dengan banyak plugin dan komunitas besar. • Mampu menghasilkan gambar berkualitas tinggi dan detail. 	<ul style="list-style-type: none"> • Sumber daya komputasi yang besar diperlukan untuk hasil optimal. • Tidak ramah pengguna untuk non-teknisi.
Chat GPT DALL-E	<ul style="list-style-type: none"> • Integrasi langsung dengan ChatGPT memudahkan pembuatan gambar dalam percakapan. • Hasil yang inovatif dan bisa disesuaikan dengan <i>prompt</i> spesifik. • Gratis dan mudah diakses melalui platform OpenAI. 	<ul style="list-style-type: none"> • Kualitas gambar kadang-kadang tidak sebanding dengan alat khusus gambar lainnya. • Kontrol kreatif terbatas dibandingkan dengan generator gambar yang lebih canggih. • Terbatas dalam gaya dan kompleksitas output.
Adobe Firefly	<ul style="list-style-type: none"> • Terintegrasi dengan ekosistem Adobe, ideal untuk desainer yang sudah menggunakan alat Adobe lainnya. • Fitur khusus untuk manipulasi gambar dan desain grafis tingkat lanjut. • Fokus pada hasil yang sesuai untuk produksi profesional. 	<ul style="list-style-type: none"> • Akses berbayar dan memerlukan langganan Adobe. • Kompleksitas tinggi untuk pemula. • Fitur terbatas untuk pengguna di luar ekosistem Adobe.
Midjourney	<ul style="list-style-type: none"> • Dikenal untuk menghasilkan gambar dengan kualitas artistik yang sangat tinggi dan unik. • Berfokus pada komunitas kreatif dengan banyak fitur kustomisasi. • Dapat menghasilkan gaya visual yang sangat beragam dan kreatif. 	<ul style="list-style-type: none"> • Akses penuh berbayar dan mahal untuk beberapa pengguna. • Memerlukan penggunaan melalui Discord, yang bisa membingungkan bagi pengguna baru. • Prosesnya bisa memakan waktu dan hasil tidak sesuai harapan jika <i>prompt</i> kurang jelas.

Pemilihan Bing *Image Creator* sebagai objek studi utama dalam penelitian *prompt engineer* ini karena beberapa alasan kunci. Pertama, Bing Image Creator terintegrasi dengan ekosistem Microsoft yang luas, seperti Bing dan Microsoft Designer, sehingga memberikan kemudahan akses serta potensi penggunaan yang lebih luas bagi pengguna di berbagai aplikasi (Feng, 2023). Kedua, platform ini dikenal mampu menghasilkan gambar berkualitas tinggi secara konsisten, terutama ketika diberikan *prompt* yang sederhana dan jelas, memungkinkan evaluasi mendalam terhadap respons AI terhadap berbagai tingkat kompleksitas *prompt* (Lee, 2023).

Selain itu, antarmuka Bing Image Creator yang *user-friendly* menjadikannya lebih terjangkau bagi pengguna umum, yang sejalan dengan tujuan penelitian untuk membantu pengguna memahami dan meningkatkan keterampilan dalam membuat *prompt*. Fleksibilitas platform ini dalam menangkap berbagai elemen visual, termasuk berbagai gaya seperti 3D Pixar, memungkinkan eksplorasi yang lebih luas dalam berbagai konteks visual. Akhirnya, dukungan komunitas dan pengembangan berkelanjutan dari Microsoft memberikan keunggulan dalam hal pembaruan fitur dan inovasi teknologi, menjadikan Bing *Image Creator* pilihan yang representatif dan relevan untuk penelitian ini.

2. Pengembangan *Prompt Engineering* Berbasis Elemen ANV

Pembuatan dan pengujian *prompt engineering* elemen ANV (*Adjective + Noun + Verb*) dilakukan dengan menggunakan *prompt* sederhana yang kemudian dievaluasi secara progresif. Setiap elemen ditambahkan secara bertahap untuk mengamati bagaimana AI menangkap dan merepresentasikan deskripsi visual yang diberikan. Proses ini tidak hanya membantu dalam mengidentifikasi kemampuan AI dalam merespons variasi *prompt* yang semakin kompleks, tetapi juga memberikan wawasan tentang bagaimana kombinasi elemen-elemen ini dapat dioptimalkan untuk menghasilkan gambar yang lebih akurat dan sesuai dengan harapan pengguna.

Tabel 3. *Prompt Engineering 1* Menggunakan AI *Image Generator*.
Sumber: Darmawan, 2024.

<p><i>Prompt Engineering 1</i> An asian glasses woman, jumping with smile and raise both hands.</p>	
<p>A : Asian, Glasses N : Woman V : Jumping, Smile, Raise both hands</p>	

Gambar yang dihasilkan dengan *prompt engineering 1* (*An asian glasses woman, jumping with smile and raise both hands.*), menunjukkan interpretasi yang baik dari elemen-elemen *prompt*. Wanita yang ditampilkan memiliki penampilan Asia dengan kacamata, sesuai dengan deskripsi "Asian" dan "Glasses" pada semua gambar. Aksi melompat, tersenyum, dan mengangkat kedua tangan terwakili dengan jelas di semua gambar kecuali gambar kiri bawah, menggambarkan dinamika dan ekspresi kebahagiaan yang diinginkan. AI berhasil menangkap gerakan dan emosi dari prompt, menciptakan gambar yang energik dan konsisten dengan deskripsi, meskipun terdapat variasi dalam gaya pakaian dan latar belakang yang tetap relevan dan memperkaya visual keseluruhan.

Tabel 4. *Prompt Engineering 2* Menggunakan AI Image Generator.
 Sumber: Darmawan, 2024.

<p><i>Prompt Engineering 2</i> <i>An asian glasses woman, wearing tank top, jumping with smile and raise both hands.</i></p>	
<p>A1 : Asian, Glasses N : Woman A2 : Wearing tank top V : Jumping, Smile, Raise both hands</p>	

Gambar yang dihasilkan dengan *prompt engineering 2* (*An asian glasses woman, wearing tank top, jumping with smile and raise both hands.*), secara keseluruhan menunjukkan pemahaman yang baik dari AI terhadap elemen-elemen deskriptif yang diminta. Dalam setiap gambar, wanita Asia yang berkacamata dan mengenakan *tank top* tampil dengan ekspresi ceria, senyum lebar, dan gestur tangan yang mengindikasikan kegembiraan. Namun, meskipun pose "Jumping" tidak selalu terlihat eksplisit, elemen "Smile" dan "Raise both hands" dieksekusi dengan konsistensi tinggi, khususnya pada gambar kanan bawah. Variasi latar belakang tetap sederhana dan tidak mengalihkan perhatian dari subjek utama, sementara pakaian dan aksesoris yang diminta jelas terlihat di setiap gambar. Hasil ini

menunjukkan bahwa AI mampu menangkap detail-detail penting dari prompt, meskipun beberapa aspek seperti "Jumping" dapat ditafsirkan lebih fleksibel.

Tabel 5. *Prompt Engineering 3 Menggunakan AI Image Generator.*
Sumber: Darmawan, 2024.

<p><i>Prompt Engineering 3</i> <i>An asian glasses woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, in Gym.</i></p>	
<p>A1 : Asian, Glasses N : Woman A2 : Hair tied, Wearing tank top V : Jumping, Smile, Raise both hands A3 : in Gym</p>	

Gambar yang dihasilkan dengan *prompt engineering 3* (*An asian glasses woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, in Gym.*), kata kunci "in Gym" secara keseluruhan menunjukkan pemahaman yang baik dari AI terhadap semua elemen yang diminta. Wanita Asia berkacamata dengan tank top terlihat di lingkungan gym pada setiap gambar, sesuai dengan deskripsi. Meskipun gerakan "Jumping" hanya tercipta pada gambar kiri atas, pose dengan tangan terangkat dan ekspresi senyum lebar tetap konsisten, mencerminkan energi dan kebahagiaan yang diinginkan. Latar belakang gym mendukung konteks aktivitas yang lebih dinamis, dengan peralatan olahraga yang terlihat jelas, memperkuat suasana tempat. Gambar ini menunjukkan bahwa AI dapat menangkap dan memadukan deskripsi visual yang kompleks dengan baik, menghasilkan gambar yang sesuai dengan *prompt* yang diberikan.

Tabel 6. *Prompt Engineering 4* Menggunakan AI Image Generator.
Sumber: Darmawan, 2024.

Prompt Engineering 4

An asian glasses woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, in Gym. And many people surprisingly looks her.

- A1 : Asian, Glasses
N : Woman
A2 : Hair tied, Wearing tank top
V : Jumping, Smile, Raise both hands
A3 : in Gym, many people surprisingly looks her



Gambar yang dihasilkan dengan *prompt engineering 4* (*An asian glasses woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, in Gym. And many people surprisingly looks her.*), menunjukkan sebagian besar elemen yang diminta dengan baik, namun ada beberapa keterbatasan. Wanita Asia berkacamata dengan rambut diikat dan mengenakan tank top terlihat dengan jelas di setiap gambar, sesuai dengan deskripsi. Ekspresi senyum lebar, gestur tangan terangkat, dan lingkungan gym juga ditampilkan secara konsisten. Namun, elemen "Jumping" tidak terwakili semua gambar, dan bagian penting dari prompt mengenai "Many people surprisingly looks her" tidak ditampilkan pada semua gambar. AI tampaknya fokus pada subjek utama, tetapi tidak menangkap interaksi sosial atau reaksi dari orang lain di sekitar, yang seharusnya menjadi elemen signifikan dalam *prompt* ini. Ini menunjukkan bahwa AI dapat menangkap detail visual utama dengan baik, namun mungkin kesulitan dalam merepresentasikan interaksi kompleks antara subjek dan elemen latar lainnya.

Tabel 7. *Prompt Engineering 5 Menggunakan AI Image Generator.*
Sumber: Darmawan, 2024.

<p><i>Prompt Engineering 5</i> <i>An asian glasses angel woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, In the middle of town. And many people surprisingly looks at her.</i></p>	
<p>A1 : Asian, Glasses, Angel N : Woman A2 : Hair tied, Wearing tank top V : Jumping, Smile, Raise both hands A3 : in Middle town, many people surprisingly look her</p>	

Gambar yang dihasilkan dengan *prompt engineering 5* (*An asian glasses angel woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, In the middle of town. And many people surprisingly looks at her.*), menunjukkan variasi dalam penerjemahan elemen-elemen yang lebih kompleks. Jika dibandingkan dengan *prompt* sebelumnya, tambahan deskripsi seperti "Angel" dan "in Middle town" seharusnya menambahkan elemen supernatural dan lingkungan kota yang ramai ke dalam gambar. Namun, AI mungkin mengalami kesulitan dalam menggabungkan karakteristik malaikat seperti sayap atau aura dengan penampilan modern wanita yang diinginkan. Demikian pula, elemen "Many people surprisingly looks her" dan pengaturan "Middle town" mungkin tidak ditangkap dengan sempurna seperti pada semua gambar, mengingat tantangan untuk menyampaikan lingkungan yang ramai dan interaksi sosial yang spesifik. Secara keseluruhan, gambar yang dihasilkan mungkin berhasil dalam menampilkan subjek utama; wanita Asia berkacamata dengan rambut terikat dan tersenyum, namun perintah melompat juga tidak terwakili pada semua gambar; tetapi kemungkinan Bing *Image Creator* kurang dapat menangkap *prompt* malaikat dan

konteks urban yang lebih dinamis serta respons orang-orang di sekitarnya. Hal ini disadari penulis adanya perbedaan konteks realis dengan imajinatif malaikat.

3. Penambahan *Prompt Engineering* Berbasis Elemen S

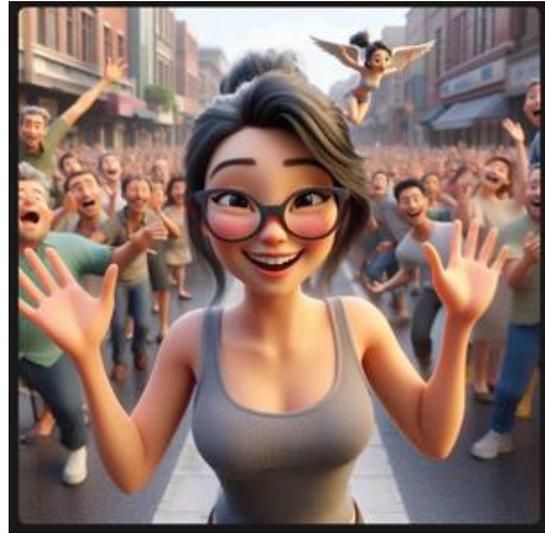
Penulis menambahkan elemen S (*Style*) dalam penelitian ini untuk menguji bagaimana AI Image Generator, seperti Bing Image Creator, menangkap dan merepresentasikan gaya visual tertentu yang diinginkan oleh pengguna. Dengan menambahkan elemen ini, penulis dapat mengevaluasi apakah AI mampu mempertahankan konsistensi gaya (seperti 3D Pixar style) sambil tetap mengakomodasi elemen-elemen visual lain yang sudah ada dalam prompt. Hal ini penting karena gaya visual mempengaruhi cara elemen-elemen lain, seperti karakteristik fisik dan konteks, direpresentasikan dalam gambar akhir. Dengan demikian, penambahan elemen S memungkinkan penelitian ini untuk mengeksplorasi dimensi artistik dan estetika dalam proses pembuatan gambar, serta mengukur kemampuan AI dalam menghasilkan gambar yang tidak hanya akurat, tetapi juga secara visual menarik sesuai dengan preferensi gaya yang ditentukan.

Tabel 8. *Prompt Engineering 6* Menggunakan AI Image Generator.
Sumber: Darmawan, 2024.

Prompt Engineering 6
An asian glasses angel woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, In the middle of town. And many people surprisingly look at her. In 3D Pixar style.



A1 : *Asian, Glasses, Angel*
 N : *Woman*
 A2 : *Hair tied, Wearing tank top*
 V : *Jumping, Smile, Raise both hands*
 A3 : *in Middle town, many people surprisingly look her*
 S : *in 3D Pixar style*



Gambar yang dihasilkan dengan *prompt engineering 6* (*An asian glasses angel woman, with hair tied up, wearing tank top, jumping with smile and raise both hands, In the middle of town. And many people surprisingly looks at her. In 3D Pixar style.*), dengan tambahan "*3D Pixar style*" memberikan kompleksitas yang lebih tinggi dalam visualisasi. Dalam gaya 3D Pixar, AI seharusnya mampu menangkap karakteristik khas seperti proporsi tubuh yang lebih ekspresif, warna-warna cerah, dan detail yang halus. Namun, meskipun elemen utama seperti wanita Asia berkacamata dengan rambut terikat, mengenakan tank top, dan tersenyum sambil melompat mungkin direpresentasikan dengan baik dalam gaya Pixar, elemen "*Angel*" seperti sayap atau aura bisa saja tidak ditampilkan dengan jelas, mengingat kesulitan AI dalam menggabungkan aspek supernatural dengan gaya yang ditentukan. Selain itu, konteks "*in Middle town*" dan "*Many people surprisingly looks her*" tetap tidak terwakili pada semua gambar, perintah "*Jumping*" tidak ada pada semua gambar, dan terutama dalam perintah menampilkan interaksi sosial dan lingkungan kota yang hidup tidak terwakili pada semua gambar. Gaya 3D Pixar mungkin berhasil menambahkan daya tarik visual, tetapi AI masih menghadapi tantangan dalam menangkap semua detail kompleks yang diminta dalam prompt ini, terutama elemen lingkungan dan reaksi sosial.

4. Analisis dan Hipotesa *Prompt Engineering* Bing Image Creator

Proses pengembangan prompt menunjukkan bahwa semakin kompleks dan spesifik deskripsi yang diberikan, semakin besar tantangan bagi AI dalam menangkap semua elemen dengan akurasi tinggi. Awalnya, deskripsi sederhana seperti atribut fisik dan gerakan (misalnya, wanita Asia berkacamata, tersenyum, dan melompat) ditangkap dengan baik. Namun, saat elemen-elemen tambahan seperti karakteristik supernatural (seperti "*Angel*") dan konteks lingkungan sosial (seperti "*in Middle town, Many people surprisingly looks her*") ditambahkan, AI mulai menunjukkan keterbatasan dalam menggabungkan semua aspek tersebut secara konsisten. Hal ini menunjukkan bahwa pemahaman AI terhadap prompt bersifat *iterative* (berdasarkan pengelompokan kata), yang terlihat dari elemen yang lebih sederhana diproses lebih baik daripada deskripsi kompleks yang memerlukan interaksi dinamis.

Bing Image Creator mampu menangkap deskripsi visual yang bersifat langsung dan konkret dengan baik, seperti ciri-ciri fisik, pakaian, dan gerakan dasar. Namun, AI tampaknya lebih kesulitan ketika harus memproses deskripsi yang melibatkan elemen abstrak atau interaksi sosial yang kompleks, seperti "*Angel*" atau "*Many people surprisingly looks her.*" Meskipun gaya visual (seperti "*3D Pixar style*") ditangani dengan cukup baik, dengan karakter yang tampil lebih hidup dan berwarna, AI sering kali gagal dalam menangkap konteks lingkungan secara penuh dan merepresentasikan reaksi sosial atau elemen supernatural yang diminta. Hal ini menunjukkan adanya bias AI untuk fokus pada subjek utama sementara mengabaikan beberapa konteks tambahan lain yang lebih luas.

Tabel 9. Analisis dan Hipotesa Prompt Engineering Bing Image Creator.
Sumber: Hasil Analisa Penulis.

Aspek	Deskripsi
Proses Pengembangan Prompt	Semakin kompleks prompt, semakin sulit AI menangkap semua elemen. Deskripsi sederhana (atribut fisik dan gerakan) ditangkap dengan baik, sementara elemen kompleks seperti karakteristik supernatural dan interaksi sosial tidak selalu berhasil direpresentasikan.
Cara Bing Image Creator Menangkap Deskripsi	Bing Image Creator efektif dalam menangkap deskripsi yang konkret seperti fisik, pakaian, dan gerakan. Namun, AI sering gagal dalam merepresentasikan konteks yang lebih luas, interaksi sosial, dan elemen abstrak seperti "Angel" dan "Many people surprisingly looks her".
Regulasi atau Kendala yang Membatasi Hasil	Kemampuan AI mungkin dibatasi oleh algoritma yang memprioritaskan detail mudah dipahami dan norma-norma sosial atau estetika tertentu. Kendala ini bisa menyebabkan AI menyederhanakan atau mengabaikan elemen-elemen abstrak atau kompleks dalam prompt.

Ada beberapa regulasi dan kendala yang tampaknya mempengaruhi kemampuan Bing Image Creator dalam menghasilkan gambar. Pertama, kemampuan AI *image generator* untuk memproses dan menggabungkan elemen-elemen yang lebih abstrak atau kompleks mungkin dibatasi oleh algoritma yang lebih fokus pada detail-detail yang bersifat lebih mudah dipahami, seperti atribut fisik atau gaya visual. Selain itu, kedua; ada kemungkinan batasan dalam hal representasi konten yang berkaitan dengan norma-norma sosial atau estetika tertentu, seperti interpretasi elemen supernatural ("Angel") atau interaksi sosial dalam ruang publik, yang mungkin dikendalikan oleh filter atau bias bawaan. Kendala ini bisa menyebabkan Bing Image Creator mengabaikan atau menyederhanakan elemen-elemen tersebut, sehingga hasil gambar tidak sepenuhnya sesuai dengan deskripsi yang diminta pengguna dalam *prompt engineering*.

KESIMPULAN

Penelitian ini menunjukkan bahwa pemahaman dan penguasaan dalam merancang prompt adalah kunci untuk memaksimalkan hasil dari AI Image Generators seperti Bing Image Creator. Melalui penerapan metode ANVS (*Adjective + Noun + Verb + Style*), dengan evaluasi progresif dari *prompt engineering* sederhana hingga yang

lebih kompleks, terbukti bahwa AI *image generator* memiliki kemampuan untuk menangkap deskripsi visual dengan baik ketika elemen-elemen tersebut jelas dan konkret. Namun, ketika deskripsi menjadi lebih abstrak atau melibatkan interaksi sosial dan konteks lingkungan yang kompleks, AI *image generator* cenderung menunjukkan keterbatasan dalam menggabungkan semua elemen tersebut dengan akurat.

Selain itu, penelitian ini menggarisbawahi pentingnya pemilihan platform yang tepat berdasarkan kelebihan dan kekurangannya. Bing Image Creator dipilih karena kemampuannya dalam menghasilkan gambar berkualitas tinggi secara konsisten, antarmuka yang ramah pengguna, dan dukungan teknologi dari ekosistem Microsoft. Meskipun demikian, beberapa kendala AI *image generator* ini juga masih mengganggu, seperti interpretasi elemen abstrak dan penggabungan konteks sosial menunjukkan bahwa ada ruang untuk pengembangan lebih lanjut. Hasil dari penelitian ini memberikan wawasan praktis yang dapat membantu pengguna dalam merancang *prompt engineering* yang lebih efektif dan sesuai dengan kebutuhan visual pengguna.

DAFTAR PUSTAKA

- Al Saud, D. S., & Alfarani, L. (2024). The Effectiveness Of Using Microsoft Bing Image Creator In Enhancing Students' Painting Performance, Painting Creative Ideas, And Attitudes Toward It. *Journal of Research Administration*, 6(1), 85-97.
- Allingham, J. U., Ren, J., Dusenberry, M. W., Gu, X., Cui, Y., Tran, D., ... & Lakshminarayanan, B. (2023, July). A Simple Zero-Shot Prompt Weighting Technique To Improve Prompt Ensembling In Text-Image Models. *In International Conference on Machine Learning PMLR*, 1(1), 547-568.
- Beyan, E. V. P., & Rossy, A. G. C. (2023). A Review Of AI Image Generator: Influences, Challenges, And Future Prospects For Architectural Field. *Journal of Artificial Intelligence in Architecture*, 2(1), 53-65.
- Chiou, L. Y., Hung, P. K., Liang, R. H., & Wang, C. T. (2023, July). Designing With AI: An Exploration Of Co-Ideation With Image Generators. *In Proceedings Of The 2023 ACM Designing Interactive Systems Conference*, 1(1), 1941-1954.
- Crowson, K., Pavlichenko, S., & Ustalov, D. (2022). A Taxonomy of Prompt Modifiers for Text-To-Image Generation. *arXiv Preprint*, 2(1), 53-65.

- Darmawan, A. J., Utomo, S. R. H., Yoo, P. M. H., Losung, A. S., Satianegara, A. C., & Oranto, R. Z. (2024). Pelatihan Komprehensif dan Memberdayakan Media Sosial pada CGA. *Jurnal Lentera Widya*, 5(2), 111-118.
- Feng, Y., Wang, X., Wong, K. K., Wang, S., Lu, Y., Zhu, M., ... & Chen, W. (2023). Promptmagician: Interactive Prompt Engineering For Text-To-Image Creation. *IEEE Transactions on Visualization and Computer Graphics*, 5(1), 63-78.
- Hao, Y., Chi, Z., Dong, L., & Wei, F. (2023). Optimizing prompts for text-to-image generation. *NeurIPS 2023*.
- Hermanto, Y. A. L. (2024). Artificial Intelligence-Image Generator (AI-IG): Characteristics, And Influence In Industrial Development Of Design Product. In *Inclusive, Sustainable, and Transformational Education in Arts and Literature*, Routledge, 1(1), 55-61.
- Lee, Y. H., & Chiu, C. Y. (2023, July). The Impact of AI Text-to-Image Generator on Product Styling Design. In International Conference on Human-Computer Interaction, Cham: Springer Nature Switzerland, 1(1), 502-515.
- Liu, V., & Chilton, L. B. (2022, April). Design Guidelines For Prompt Engineering Text-To-Image Generative Models. In *Proceedings Of The 2022 CHI Conference On Human Factors In Computing Systems*, 1(1), 1-23.
- McGee, R. W. (2023). Comparing Gab's AI Image Generator to Microsoft Bing's Image Maker: An Experimental Study, *SSRN*, 1(1), 216-232.
- McGee, R. W. (2023). Sex, Politics and Artificial Intelligence: A Comparison of the Gabby and Bing Image Creators, *SSRN*, 1(1), 156-167.
- McGee, R. W. (2023). Using ChatGPT and Bing Image Creator to Create Images of Martial Artists: An Application of Artificial Intelligence to Create Art, *SSRN*, 1(1), 466-526.
- Noel, G. P. (2024). Evaluating AI-Powered Text-To-Image Generators For Anatomical Illustration: A Comparative Study. *Anatomical Sciences Education*, 17(5), 979-983.
- Oppenlaender, J. (2023). Prompt design and interpretation in text-to-image models: Insights from autoethnographic research. *Journal of Creative AI*, 8(2), 123-140.
- Schuhmann, C., Beaumont, R., & Vencovská, J. (2023). BeautifulPrompt: Towards automatic prompt engineering for text-to-image synthesis. *arXiv preprint*.
- Wang, M., & Lin, J. (2023). Exploring user-oriented prompt engineering in Bing AI generative tools. *Journal of AI Research Applications*, 14(3), 87-102.
- Xu, Z., Liu, T., & Zhang, W. (2023). Addressing aesthetic and semantic gaps in AI-generated images. *Journal of AI and Creativity*, 12(1), 45-59.