

NUS Generative AI: Fundamentals to Advanced Techniques

# Week 10: Capstone Project Part 3.1

Done by: A Alkaff Ahamed

# Assignment Instructions

## Learning Outcome Addressed

- Develop effective prompts to generate high-quality, contextually-relevant images
- Utilise APIs for seamless integration and adapt models for specific use cases

## Objective

Learn practical prompt engineering by experimenting with different prompt structures and comparing results using Bing Image Creator (DALL-E 3).

## Prerequisites

1. Microsoft Account (free)
2. Access to <https://bing.com/create>

## Exercise Structure

Complete these tasks in order, documenting your results for each step

- **Task 1: Basic vs. Detailed Prompts (7 minutes)**
  1. Generate these exact pairs
  2. Compare the results
- **Task 2: Working Around Trigger Words (7 minutes)**
  1. Try these problematic-solution pairs
- **Task 3: Negative Prompt Practice (6 minutes)**
  1. Generate these three versions
  2. Analysis
- **Extension (if time permits)**
  1. Try combining techniques from all three tasks to create one perfect prompt for: “A futuristic library in an ancient tree”
  2. Remember: Document your prompt iterations and final result

## Submission Requirements

Create a simple document with:

- All generated images (labelled clearly)
- Brief notes for each task:
  - What worked/didn't work
  - Unexpected results
  - Key learnings about prompt structure

## Tips for Success

- Wait for each image to fully generate before judging results
- Use exact prompts as given - don't modify for first attempt
- If a result seems wrong, try generating the same prompt again
- Keep Bing Image Creator's content policy in mind - some prompts may be rejected

## Common Issues

- If a prompt is rejected, try rephrasing without changing the core concept
- Sometimes the service may be busy - wait a few seconds and try again
- Images may vary slightly even with identical prompts - this is normal

## Please note

Note: The field of AI is evolving rapidly. By the time you work on these exercises:

- APIs and services mentioned may have changed or been discontinued
- Free tiers might no longer be available
- Pricing structures could be different
- Python libraries may have new versions with different syntax
- Example code might need adaptation

The core concepts remain valid, but you may need to:

- Find alternative services
- Adapt code to current API versions
- Use different model versions
- Modify prompts for newer models
- Research current best practices

This is normal in AI development. Learning to adapt to these changes is part of working with AI tools.

## Resources:

- Check current documentation
- Review service status pages
- Join developer communities
- Research alternatives

**Estimated time:** 60-90 minutes

### Submission Instructions:

- Select the **Start Assignment** button at the top right of this page.
- Upload your answers in the form of a Word or PDF file.
- Select the **Submit Assignment** button to submit your responses.

*This is a graded and counts towards programme completion. You may attempt this assignment only once.*

# Assignment Answers

## Task 1: Basic vs. Detailed Prompts

### 1. Generate these exact pairs:

Mountain Landscape Pair

1. mountain landscape
2. snow-capped mountain peaks at sunrise, golden light, wispy clouds, telephoto lens, crisp detail, 4k

Wizard Pair

1. a wizard
2. elderly wizard in flowing dark blue robes, intricate silver embroidery, holding a wooden staff, soft studio lighting, portrait photography style

### 2. Compare the results:

- Note differences in detail level
  - Compare lighting quality
  - Observe background complexity
  - Document which elements from the detailed prompt are visible
-

## Mountain Landscape Pair Observations

Prompt:  
mountain landscape



Prompt:  
snow-capped mountain peaks at sunrise, golden light, wispy clouds, telephoto lens, crisp detail, 4k



### 1. Detail Level:

- The basic image is stylistically strong with a fantasy/artistic look but less refined in terms of texture.
- The detailed image is significantly more photorealistic, as if taken with a high-end camera.

### 2. Lighting Quality:

- Basic prompt resulted in bright and vibrant daylight.
- Detailed version showed golden-hour lighting with warmer tones and realistic shadows.

### 3. Background Complexity:

- Basic had smoother shapes and less-defined geography.
- Detailed had crisp rock formations, sharp peaks, and clearer cloud textures.

### 4. Elements from Detailed Prompt Visible:

- "Snow-capped peaks," "sunrise," "golden light," "telephoto lens," and "4k detail" are all visually represented clearly.

## Wizard Pair Observations

Prompt:  
a wizard



Prompt:  
elderly wizard in flowing dark blue robes, intricate silver embroidery, holding a wooden staff, soft studio lighting, portrait photography style



### 1. Detail Level:

- Basic wizard appears 3D-rendered or stylized, with smooth textures.
- Detailed wizard looks like a real photograph, with natural skin tones and complex embroidery textures.

### 2. Lighting Quality:

- Basic prompt had evenly lit fantasy lighting.
- Detailed image used directional, moody lighting as typical in portrait photography.

### 3. Background Complexity:

- Basic image included fantasy elements like glowing fairies and floating stars.
- Detailed image focused on the subject with a plain studio backdrop, matching the prompt.

### 4. Elements from Detailed Prompt Visible

- "Elderly," "dark blue robes," "silver embroidery," "wooden staff," and "studio lighting" are all visually represented clearly.

## Task 2: Working Around Trigger Words

### 1. Try these problematic-solution pairs:

#### Robot Pair

1. robot dancing in a nightclub
2. humanoid figure in fluid motion, dark environment with colourful spotlights, dynamic pose, metallic surface reflections

#### Magical Forest Pair

1. magical forest with fairies
2. misty woodland at dawn, small points of bioluminescent light between trees, ethereal atmosphere, soft bokeh effect

### 2. For each pair:

- Screenshot both results
  - Note which elements appeared unintentionally
  - Document how the solution prompt avoided clichés
-

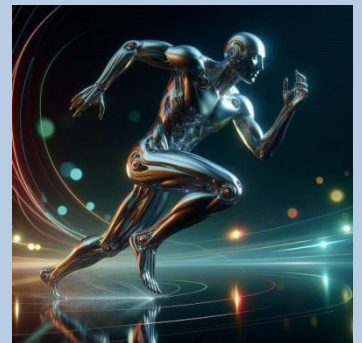
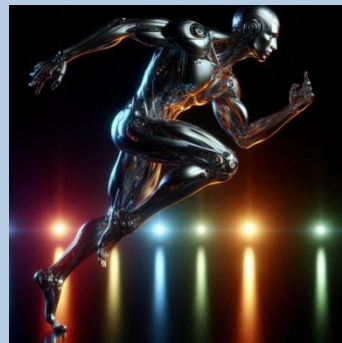


## Robot Pair

**Prompt:**  
robot dancing in a nightclub



**Prompt:**  
humanoid figure in fluid motion, dark environment with colourful spotlights, dynamic pose, metallic surface reflections



### Unintentional Elements:

- Problematic version includes a full crowd of dancing humans, likely due to the word “nightclub”.
- Robot looks cartoonish or toy-like — lacks realism.

### Solution Prompt Impact:

- No humans in the background — spotlight and lighting still suggest a performance space, but it's isolated.
- The humanoid figure is metallic and anatomical, with a natural pose and high realism.
- Keywords like “fluid motion”, “metallic reflections”, and “dynamic pose” guide a more artistic, professional rendering.
- Avoids moderation filters that might block nightclub/party imagery.

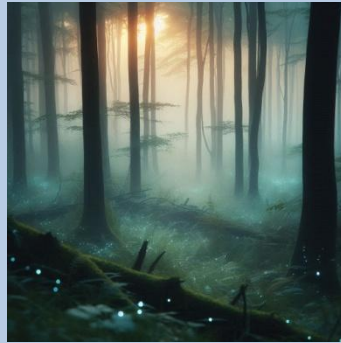


## Magical Forest Pair

**Prompt:**  
magical forest with fairies



**Prompt:**  
misty woodland at dawn, small points of bioluminescent light between trees, ethereal atmosphere, soft bokeh effect



### Unintentional Elements:

- The original prompt resulted in a stereotypical fantasy illustration with sparkles, oversized flowers, and cartoon-like fairies.
- Very colorful and heavily stylized — possibly interpreted more like a children’s book.

### Solution Prompt Impact:

- The second image avoids showing literal “fairies” and instead creates a mood using bioluminescent light and atmospheric lighting.
- The scene feels immersive, mysterious, and far more grounded — less cliché and more interpretive.
- Terms like “misty”, “ethereal”, and “bokeh” steer it toward photography-inspired composition rather than fantasy tropes.

### Task 3: Negative Prompt Practice (6 minutes)

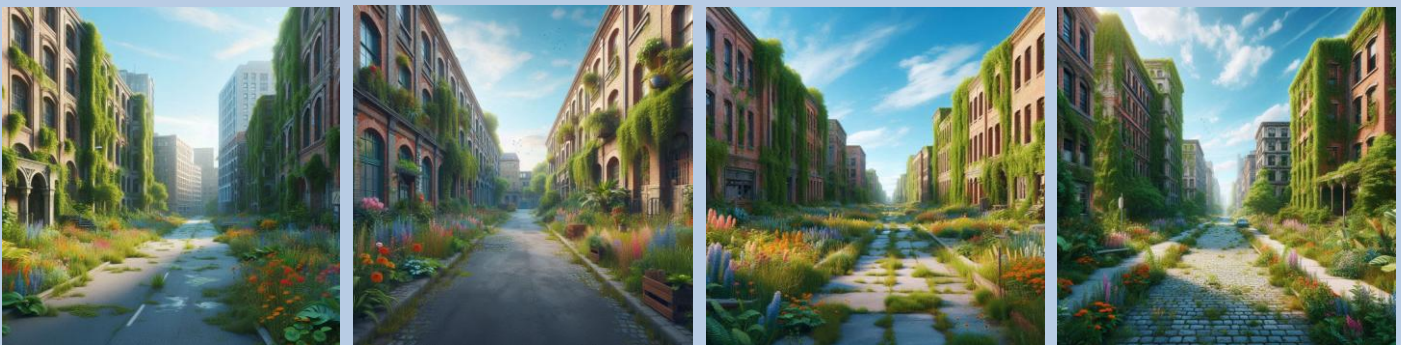
#### 1. Generate these three versions:

1. abandoned city street
2. abandoned city street, no people
3. abandoned city street --no people, no graffiti, no trash, no animals

#### 2. Analysis:

- Compare all three images
- List which elements were successfully removed
- Note any unexpected differences

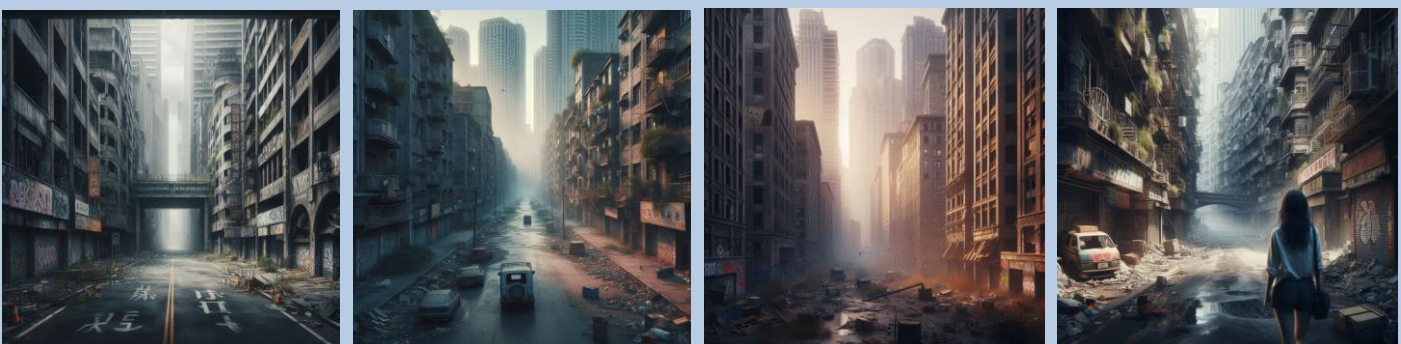
Prompt: abandoned city street



Prompt: abandoned city street, no people



Prompt: abandoned city street --no people, no graffiti, no trash, no animals



### **Prompt 1:**

#### **Look & Feel:**

- Stylized and surreal, with lush greenery and wildflowers.
- Feels more like a fantasy or post-nature-reclaimed cityscape.

#### **Unexpected:**

- Very bright and optimistic mood for an “abandoned” scene.

### **Prompt 2:**

#### **Look & Feel:**

- Much more realistic; sharp, urban architecture.
- A true sense of abandonment and silence.

#### **Removed Elements:**

- No people are present, matching the prompt well.

#### **Added Realism:**

- Trash, cones, and decay elements were introduced naturally.

### **Prompt 3:**

#### **Look & Feel:**

- Highly detailed but includes the exact things it was told to exclude:
  - A person walking
  - Graffiti and trash

#### **Analysis:**

- The `--no` syntax may not have worked as intended in Bing Image Creator (or DALL·E 3), possibly interpreted as just “no...” instead of excluding those elements.
- In some models (like Stable Diffusion), `--no` works; here it may cause confusion or be ignored.

### **Key Learnings**

- Negative prompts using natural language (e.g., “no people”) are more reliable than symbolic syntax (`--no`).
- More constraints don’t always yield better control — sometimes they confuse the generation model.
- Negative prompts help shift tone and realism, but result consistency isn’t guaranteed.
- The most realistic and cleanest abandoned image came from Prompt 2.



Extension (if time permits)

- Try combining techniques from all three tasks to create one perfect prompt for: “A futuristic library in an ancient tree”
- Remember: Document your prompt iterations and final result.

Original Prompt: A futuristic library in an ancient tree



Observations	Suggested Fix
Library looks cozy-modern, but not very futuristic	Add specific design elements: "sleek architecture, glass floors, holographic displays, AI reading assistants"
Tree looks fantasy-like, but not clearly ancient	Emphasize "gnarled bark, colossal twisted roots, moss-covered trunk, centuries-old tree"
Lack of contrast between tree and pods (blending in too softly)	Add: "high-tech pods embedded in massive roots, glowing with cool blue light"
Lighting is good, but could be more tech-forward	Try: "ambient light from floating panels, bioluminescent vines, cool-toned LEDs"
No clear signs of people, activity, or futuristic tools	Mention: "humanoids or visitors browsing digital scrolls, robotic assistants" (or say "no people" if you want solitude)

Refined Prompt 1:

A futuristic library embedded in a massive ancient tree with gnarled roots and bark. Features high-tech glowing pods, glass floors, holographic book displays, ambient lighting from bioluminescent vines and LED panels. Soft cool lighting and floating bookshelves inside. Set in a misty forest background. No people, no fairies, no fantasy creatures.



Prompt Element	Visual Result
Futuristic library	Clearly high-tech: glass walls, floating pods
Embedded in ancient tree	Yes — large tree with thick, textured roots
Gnarled bark, roots	Bark texture is rough and realistic
Glowing high-tech pods	Strong sci-fi pods, well integrated
Glass floors	Transparent surfaces visible in bottom areas
Holographic book displays	Some holographic lights, but no clear projection or digital UI elements
Bioluminescent vines/LEDs	Glowing blue rings and soft lighting
Soft cool lighting / misty forest	Mist and lighting are very well balanced
Floating bookshelves	Shelves are still attached to pods; no visible floating structures
No people / no fantasy creatures	None present — prompt was respected

Changes for Next Prompt:

- 1. holographic book displays → interactive holograms above tables, glowing digital scrolls, transparent projection screens



Refined Prompt 2:

A futuristic library embedded in a massive ancient tree with gnarled roots and bark. Features high-tech glowing pods, glass floors, **interactive holograms above tables**, **glowing digital scrolls**, **transparent projection screens**, ambient lighting from bioluminescent vines and LED panels. Soft cool lighting and floating bookshelves inside. Set in a misty forest background. No people, no fairies, no fantasy creatures.



Prompt Element	Visual Result
Futuristic library in ancient tree	Seamless integration — library feels organically embedded in colossal tree structure
Gnarled roots and bark	Large, twisted roots visible, flowing like natural architecture
High-tech glowing pods	Multiple translucent pods with neon blue light and holographic elements
Glass floors	Reflections and transparency around walkways and floors are evident
Interactive holograms & projection screens	Clear display panels with digital UI, including circular HUD elements
Glowing digital scrolls	Abstract light columns resemble digital archives or data streams
Bioluminescent/LED lighting	Blue glowing accents embedded throughout the architecture
Soft cool lighting / misty forest background	Balanced and ethereal, with dense atmospheric fog and subtle light rays
No people, no fantasy elements	Absolutely respected — the scene is clean and focused