Generative AI

An Introduction to text and image generation using Generative AI

V1.0

What is Generative AI

- It is a subset of Artificial Intelligence / Machine Learning
  - Uses text input to generate text or image output

Why should photographers care about it?

- A new way of creating strong images
  - Can be incorporated into our images in post to create something stronger than we were able to using just the camera
- May simplify other tasks related to photography
  - Artist's statements
  - Image titles

What tools will we be looking at

- ChatGPT – text generator
- Image Generators
  - Dall-E - https://platform.openai.com/apps
  - Stable Diffusion - https://stablediffusionweb.com/
  - **Midjourney** - https://www.midjourney.com/home/
- Photo Editing
  - Content Aware Functions in current Adobe Products
    - Uses image data, rather than text data
  - Adobe Firefly (in Photoshop) – Generative Fill

NOTE:

- These tools require a fast internet link, something we do not have at the Kanata Senior's Centre
  - Images and text are generated by powerful online computers by the software supplier
  - Will be showing screen captures of these processes.

Legal / Ethical Issues

- As these tools are quite new, the ownership of generated content and copyright status are not well defined
  - Recent rulings in the USA suggest that the works cannot be copyrighted as they have not been produced by humans.
    - These can be turned into copyrightable works through human effort
  - Content producers are challenging the legality of the training sets used
  - The actors / screenwriters strike is largely about use of this technology in the film and television industries
- These are NOT photographs and should not be passed off as being one
  - These are computer generated images

Access

- Many of these tools are currently available **on line** to try
  - There are "free" and "Paid" options
  - **Must sign up for an account**
- Adobe products have some of this technology built into recent releases
  - Firefly is now available in various Adobe products

ChatGPT –Text Generation

- Product from Open AI
  - Microsoft is a major shareholder and is integrating software into Edge browser
  - Version 3.5 is free
  - Version 4.0 requires paid subscription
  - https://www.openai.com

USE OF CHATGPT

- I HATE WRITING "ARTIST'S STATEMENTS", SO WILL USE CHATGPT TO GENERATE ONE
  - WILL LIKELY NEED A NUMBER OF ITERATIONS WITH THE SOFTWARE TO GET AN ACCEPTABLE OUTPUT

GPT-3.5 GPT-4

ChatGPT

Write a SQL query
that adds a "status" column to an "orders" table

Come up with concepts
for a retro-style arcade game

Explain nostalgia
to a kindergartener

Create a content calendar
for a TikTok account

Send a message

Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts. ChatGPT August 3 Version

INPUT TO CHATGPT:

create a 2 paragraph artist's statement about photographing glass that has water in it and explores the reflection and refraction of light in the scene. Add that with his engineering background, the photographer uses elements of orthographic projection in his work. The work often incorporates colored lights and smoke to get the desired effect.

ChatGPT Output = 100% Fluff

My photographic odyssey embarks on a voyage through the captivating interplay of glass, water, and the ethereal nature of light. Through my lens, I unveil the mesmerizing dance of reflection and refraction, transforming mundane scenes into extraordinary visual symphonies. With a foundation in engineering, I interlace elements of orthographic projection into my work, crafting images that bridge precision and artistry, offering a unique perspective on the convergence of form and illumination.

Each image I create is an invitation to explore the intricate dialogue between glass and water, where surfaces become canvases for light's intricate brushstrokes. By integrating my engineering acumen, I capture these moments with precision, unveiling the hidden dimensions that underlie everyday scenes. To achieve the desired resonance, I merge colored lights and delicate plumes of smoke, elevating the effects of reflection and refraction to ethereal heights. This fusion of technical prowess and artistic vision culminates in imagery that both tantalizes the senses and encourages introspection. My goal is to transport viewers into the realm where the tangible and the intangible converge, forging a connection between the rational and the sublime.

CHANGE INPUT TO CHATGPT:

create a 2 paragraph artist's statement, **with minimal flowery language,** about photographing glass that has water in it and explores the reflection and refraction of light in the scene. Add that with his engineering background, the photographer uses elements of orthographic projection in his work. The work often incorporates colored lights and smoke to get the desired effect.

NEW CHAT GPT OUTPUT

In my photographic endeavors, I focus on the captivating interplay between light, glass, and water. My lens delves deep into the realm of reflections and refractions found within transparent surfaces. Drawing from my engineering background, I employ elements of orthographic projection to ensure a precise and clear portrayal of these dynamic interactions.

My creative approach involves infusing colored lights and subtle wisps of smoke into the scenes, enhancing the visual impact. These elements blend harmoniously to create an atmosphere that brings out the desired effect. Through meticulous experimentation and thoughtful composition, I aim to capture moments where light bends, dances, and transforms within the glass and water, revealing a world of beauty and intrigue that might otherwise go unnoticed. This merging of technical expertise and artistic vision is at the heart of my exploration, resulting in imagery that evokes wonder and appreciation for the marvels of light and its interaction with the materials around us.

IMAGE GENERATION

AI GENERATED IMAGES OF A SADHU

DALL-E

Stable Diffusion

Midjourney

Image Generators

- Midjourney currently produces the most photo-realistic images and tends to do Arms, legs, hands, feet, fingers and toes better than the others
- Is accessed through DISCORD instant messaging platform
  - Midjourney is challenging to set up on discord platform
    - Once it is running, it works well
    - Output has a maximum size of 1024 x 1024 pixels at this point

Discord

manfredm's server

general

Search

TEXT CHANNELS

general

VOICE CHANNELS

General

manfredm

U1 U2 U3 U4

V1 V2 V3 V4

Midjourney Bot BOT Today at 8:01 PM

harbour in hong kong full of junks - @manfredm (fast)

U1 U2 U3 U4

V1 V2 V3 V4

Voice Connected

General / manfredm's server

manfredm

manfredm.

Message #general

8:03 PM

2023-08-21

INPUT TO MIDJOURNEY

- /IMAGINE PHOTO OF A SADHU SITTING ON STEPS OF A GHAT WITH STEPS FILLING THE BACKGROUND –STYLE RAW

MIDJOURNEY OUTPUT

After Retouch

Adobe products

- Some features are built into production versions
  - Sky replacement
  - Content aware fill / scale
  - Spot Healing Brush, etc.
- Generative AI (Firefly)
  - Uses text input

File Edit Image Layer Type Select Filter 3D View Plugins Window Help

Mode: N Type: Content-Aware Create Texture Proximity Match Sample All Layers 0°

Share

Cruet and Glasses 01a.psd @ 22.2% (Layer 6, RGB/16*) *

manfredm._photo_of_a_sadhu_sitting_on_steps_of_a_ghat_with_the__2fed19d7-bf9d-4b5a-8775-e1a655baa5b2.png @ 100% (Layer 0, RGB/8#) *

Layers

Kind

Normal Opacity: 100%

Lock: Fill: 100%

Layer 0

Properties

Pixel Layer

Transform

W 13.33 in X 0 in

H 16.67 in Y 0 in

0.00°

Align and Distribute

Align:

Quick Actions

Remove Background

Select Subject

View More

100% 13.333 in x 16.667 in (72 ppi)

9:03 AM
2023-08-22

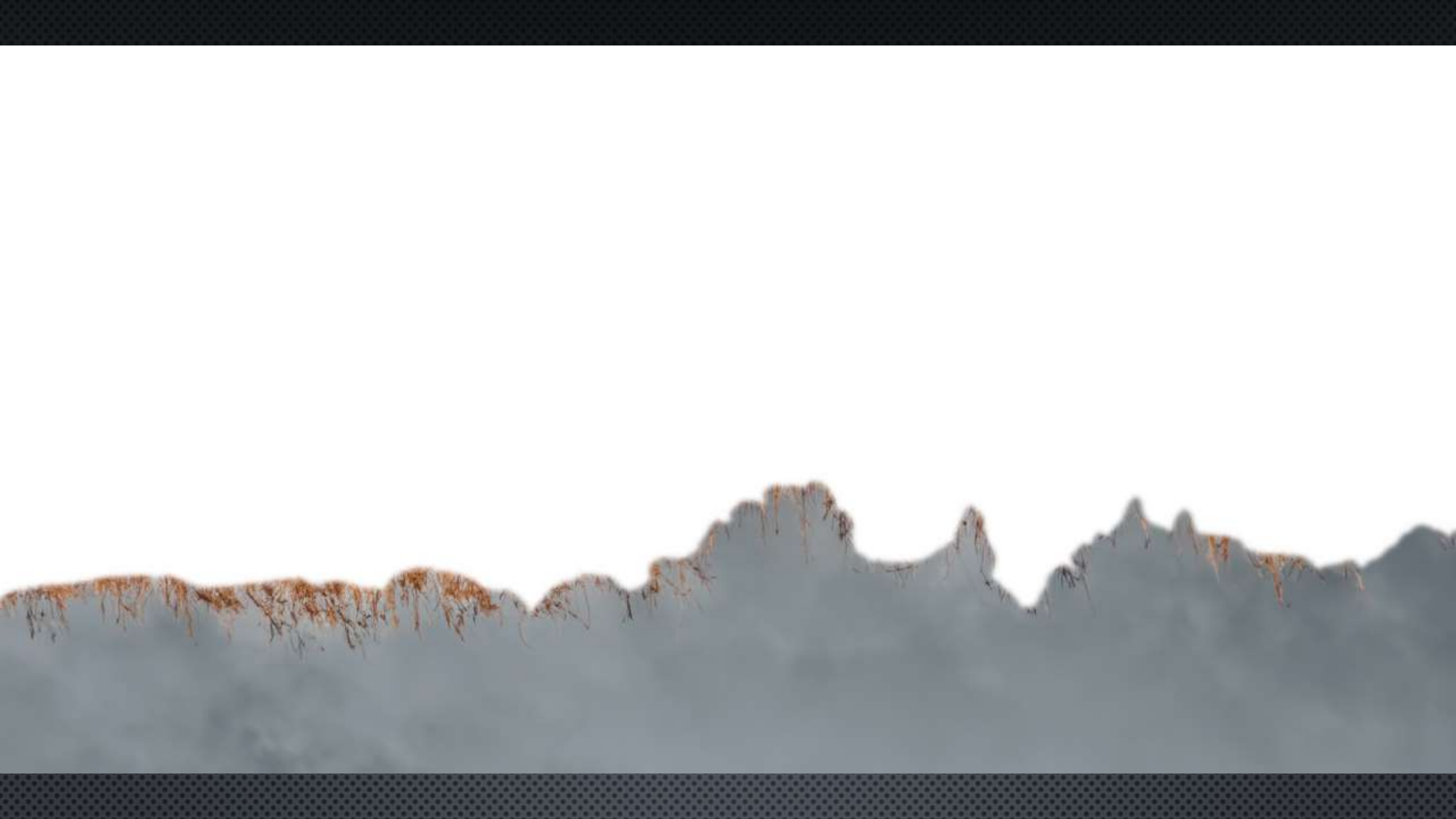
GBL
RALCO

File Edit Image Layer Type Select Filter 3D View Plugins Window Help
Sample All Layers
Enhance Edge
Select Subject
Select and Mask...
Share
Cruet and Glasses 01a.psd @ 16.7% (Hue/Saturation 1, Layer Mask/16) *
Stemless Glassware A 4x5.psd @ 16.7% (Curves 25, Layer Mask/16) *
2 boats and grasses SOOC.psd @ 25% (Layer 4, RGB/16) *
Bicycle Shop 01.psd @ 16.7% (Layer 2, RGB/16*) *
Layers
Kind
Normal
Opacity: 100%
Lock:
Fill: 100%
Layer 2
Curves 15
Curves 16
Curves 17
Curves 18
Layer 1
Properties
Pixel Layer
Transform
W 10.18 in
X 1.18 in
H 13.74 in
Y 4.06 in
0.00°
Align and Distribute
Align:
Quick Actions
Remove Background
Select Subject
View More
16.67%
11.067 in x 19.103 in (360 ppi)

GRL
RALCO

GBL
RALCO

GRL
RALCO



The future is here!

- Sort of…
  - Stay tuned!