# Assignment

## Problem Statement: *Place an object's image in a text-conditioned scene*

Recent advancement in generative AI has led to a development of a lot of creative workflows. One such workflow is to use generative AI techniques for creating realistic product photographs (traditionally done in a studio) to display a product on e-commerce websites to appeal to users. The concrete problem is given an image of an object with a white background (say Figure 1), generating a text-conditioned scene with the image placed naturally in the scene with the final output looking coherent (e.g. Figure 2).

"Realism" of the scene is an open-ended challenge and the perception of realism can depend on a lot of factors like aspect ratio of the object relative to the scene, spatial placement of the object, lighting of the scene matching the object. The task is to come up with approaches to make the scene as natural looking as possible.

You can go through the references for additional guidance and feel free to write at harsh.maheshwari@avataar.ai for any clarifications.

## Tasks:

1. The task is to write an executable code that takes the location of the image and the text prompt from the command line argument and outputs a generated image.
   (e.g. `python run.py --image ./example.jpg --text-prompt "product in a kitchen used in meal generation" --output ./generated.png`)
   You can use any libraries and publicly available trained checkpoints in achieving this goal.
   The generated image should be:
   a. Aligned with the given text prompt
   b. Should have the given object image unaltered
   c. Should be natural looking
2. Now that you have some solution for the first (hopefully interesting) problem, let's get a bit more creative. Try creating a small video output by generating multiple such consistent frames. The video can be of the camera zooming out of the scene gradually or translating in a random direction, or you can get more ambitious by moving the object in a certain direction. Go crazy!

## Examples:


Figure 1.

Figure 2. (text-prompt: "Product in a kitchen used in meal preparation")

## Attachments:

Example input images that you can try your algorithm with suitable text-prompts are attached here.

## Deliverables:

As part of the submission, provide a link of the github repository with the code and a readme file with the results on the attached images (and other examples that you may want to show). Also describe how to execute the code, how did you approach the problem and any visual results from failed and successful experiments, if any.

## Reference:

runwayml/stable-diffusion-inpainting
stabilityai/stable-diffusion-2-inpainting