This project is an analysis of text-to-image artificial intelligence generators. The comparison will mostly focus on the newly-released DALL-E 2, but will also include two other AI art producers from earlier generations. Each AI generator will be fed the same text prompt for the analysis. Three metrics will be used to analyze the generated paintings: aesthetic, comprehension and interpretation, and creativity. This project will result in a conclusion and a recommendation for the improvement of future AI art generators based on a comparison of the performance of several AI art generators and different text prompts.

### Material and Methodology

**Material:**
- All Twitter posts of DALL-E 2 as I currently have no access to DALL-E 2, the only source that can be drawn from DALL-E 2 are from Twitter. This project will therefore collect artwork created by DALL-E 2 from Twitter posts.
- The spectrum is limited to the arts and excludes photographs. In addition, I will utilize the identical text prompts to feed the other two AI art generators and evaluate the performance of the different generators by comparing their outputs.

**Methodology:**
- There is no access to the code underlying these models, thus all evaluation will be based on text input and output images.
- All of the text prompts will include an indication of a certain art style and at least one from the subject identification and activity description.
- The three metrics developed for this project are aesthetic, comprehension and interpretation, C&I, and creativity. The aesthetic will be the formal analysis of the images produced from the perspective of human art historians. Composition, color palette, and lines and shapes will be the primary factors for conducting the formal analysis. The comprehension and interpretation metric will assess the accuracy with which you comprehend and interpret the text prompt in terms of artistic style, subject matter, and iconography. The creativity will investigate the originality of combining the formal components of the particular art style with the narrative and iconography.

### Results

**Comparison with AI Art Generators from Earlier Generation:**

<table>
<thead>
<tr>
<th>Hotpot</th>
<th>Aesthetic: 6/10</th>
<th>C&amp;I: 3/10</th>
<th>Creativity: 5/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starryai</td>
<td>Aesthetic: 8/10</td>
<td>C&amp;I: 6/10</td>
<td>Creativity: 6/10</td>
</tr>
<tr>
<td>DALL-E 2</td>
<td>Aesthetic: 9/10</td>
<td>C&amp;I: 9/10</td>
<td>Creativity: 7/10</td>
</tr>
</tbody>
</table>

- Modern & pre-modern styles:
  - pre-modern style: "The dog in the style of a renaissance painting."
  - modern style: "Remembrance of nostalgia, surrealist painting by Dalí."

**Comparison with Different Text Prompts Using DALL-E 2:**

<table>
<thead>
<tr>
<th>Hotpot</th>
<th>Aesthetic: 6/10</th>
<th>C&amp;I: 3/10</th>
<th>Creativity: 5/10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starryai</td>
<td>Aesthetic: 8/10</td>
<td>C&amp;I: 6/10</td>
<td>Creativity: 6/10</td>
</tr>
<tr>
<td>DALL-E 2</td>
<td>Aesthetic: 9/10</td>
<td>C&amp;I: 9/10</td>
<td>Creativity: 7/10</td>
</tr>
</tbody>
</table>

- DALL-E 2 generates art by combining the most distinctive and recognizable features of the subject and the style. These "features" may include facial characteristics, costumes, hairstyles, makeup, accessories, color palettes, brushstrokes, modeling of light and shadow, compositions, lines and shapes, etc. But here comes the question, how does DALL-E 2 choose which features to combine? When text prompts include the name of the style for the artist's last name if the style is named after the artist, DALL-E 2 is more likely to select the formal stylistic features. In the case above, when "Vermeer" appears as a style, DALL-E 2 generates work with Vermeer's distinctive sketchy brushstrokes and bluish, cold-toned color palette. While the first does not incorporate Vermeer's painting style.

### Conclusion and Recommendation

Comparing the performance of AI art generators, DALL-E 2 outperforms earlier generations of AI art generators on all three metrics. It can generate images with a high level of aesthetic quality, an accurate interpretation of the text prompt, and some creativity in blending information with style. Nonetheless, the outcome demonstrates that DALL-E 2 has several limitations. First, its spelling ability is relatively poor. When asked to generate graphics with some text on them, typographical errors are quite probable. Several DALL-E 2 users are also aware of this shortcoming. Second, it has different levels of art style comprehension. It has a greater understanding of postmodern and contemporary art styles, especially digital art and some cartoon styles linked to popular animations. According to one of the user reports, DALL-E 2 has trouble assigning specific attributes to particular characters. This circumstance occurs when the text prompts involve two or more figures and indicate distinct characteristics for each figure. In addition to some fundamental characteristics like gender, DALL-E 2 can easily mix up age, hairstyle/color, and clothing. Even while DALL-E 2 exhibits its strength in analyzing and comprehending subjects, it cannot create satisfactory results when the text prompt contains a novel subject, as stated in the same user report.

The majority of these constraints can be overcome by modifying the parameters of the DALL-E 2 model. For example, the disparity between the amounts of accessible digital data for works of art generated throughout different eras is the primary cause of different degrees of comprehension of art styles. The majority of the artworks created throughout different time periods, processes in the DALL-E 2 training data. This bias in the training data results in various levels of art style comprehension. However, this could be improved by altering the parameter to have more pre-modern iterations than post-modern iterations.

Currently, there are numerous critiques about the ethical issues posed by Deepfakes created by AI art generators. However, as several users have pointed out, DALL-E 2 appears to have deliberate flaws in its ability to generate photo-realistic human faces. Some say that this flaw is one of DALL-E 2’s defects. However, DALL-E 2 is capable of producing photorealistic images of objects and non-human animals. Therefore, it is more plausible to believe that the flaw is an intentional attempt to prevent the creation of Deepfakes. One of the additional worries regarding DALL-E 2 is that the AI art generators may lead to the unemployment of artists, particularly digital artists. DALL-E 2’s exceptional creativity can occasionally surpass human intelligence, as it can produce combinations of style and content that have never been observed by humans. However, rather than eliminating employment, AI art producers are more likely to change them. For instance, AI art generators like DALL-E 2 require domain expertise to improve the performance.