Evaluating How Well Open-Source AI Models Interpret Written Prompts

Madeline Pysher

Follow this and additional works at: https://cedar.wwu.edu/wwu_honors

Recommended Citation
Pysher, Madeline, "Evaluating How Well Open-Source AI Models Interpret Written Prompts" (2024). WWU Honors College Senior Projects. 806.
https://cedar.wwu.edu/wwu_honors/806

This Project is brought to you for free and open access by the WWU Graduate and Undergraduate Scholarship at Western CEDAR. It has been accepted for inclusion in WWU Honors College Senior Projects by an authorized administrator of Western CEDAR. For more information, please contact westerncedar@wwu.edu.
Evaluating How Well Open-Source AI Models Interpret Written Prompts

Madeline Pysher
490 Honors Senior Capstone Project

Abstract

The purpose of this study was to take a cursory look into understanding how good “utopian” urban form is interpreted by AI. The importance of this study is that AI is now being used in every facet of society. Some examples of this include using AI to find cures for diseases (Heaven, 2023), integrating with geography to create digital-twins that control traffic lights (Digital-Twin, n.d.), and fabrication of news and profiles on social media (Mishra, 2024). All this exposure to AI feeds into people's expectations and desires for an ideal world - aka for a utopia. Some current examples of how this is occurring include the Downtown Circle in Dubai, The Line in Saudi Arabia, Telosa in the USA, and Dogen City in Japan. These projects all incorporated renderings done in part with AI image generation software. The methodology of this study included first a selection of software for generating an image using AI. A variety of AI image generation software was assessed before playground.ai was selected. Then a variety of futuristic and nostalgic thinking utopian literature was reviewed to distill criteria to be instructed to the playground.ai model. The clarified criteria were then fed to the AI model in three different variations, birds-eye view, streets-eye view, and night life view. The images were scored from zero to two on how well they met the criteria, with a total perfect score being 10. The findings from this study found that AI image generation models do better with clear, simple, and short prompts.

Table of Contents

Abstract ........................................................................................................................................... 1
Table of Contents .............................................................................................................................. 1
Introduction/Problem ...................................................................................................................... 2
Literature .......................................................................................................................................... 3
    Futuristic ...................................................................................................................................... 3
    Nostalgic ...................................................................................................................................... 4
Methods ......................................................................................................................................... 5
Introduction/Problem

The idea of a utopia from the very beginning has been an exercise of imagination. An imagination of a world where order dominated chaos (Nabil, 2024). An imagination of a world free of the burdens of economic and administrative structures (Nabil, 2024). This paper seeks to address how our understanding of good “utopian” urban form is interpreted by AI.

Understanding how AI interprets good urban form is important because the introduction of AI changes the way decisions are made. AI image generation allows the community and planners to visualize various proposed changes to the urban form while receiving feedback and applying changes in record time, facilitating a more participatory planning process (Ghisleni, 2024). Another example of how AI changes the way decisions are made is in Chattanooga, Tennessee where the combination of AI, geographic information systems, and real-time sensors create a digital-twin that can work to reduce traffic congestion (Digital-Twin, n.d.). AI is also being used in the medical field, to combat cancer, tissue samples can be exposed to different drugs and evaluated by machine-learning models to facilitate finding the best drug cocktail for that specific person (Heaven, 2023). Unfortunately, not all uses of AI are so altruistic. Some are downright insidious, such as disinformation networks. Social media algorithms can boost fake profiles that spread divisive posts intended to influence people against certain demographics (Mishra, 2024). When dealing with AI it is important to recognize what ethical limitations should be adhered to.

While AI is great at reiterating information, it is unable to extrapolate beyond the information given to it. Manufactured or incorrect responses from an AI are characterized as hallucinations. This term was chosen to extend the narrative that even in their fallibility “they are in the process
of birthing an animate intelligence on the cusp of sparking an evolutionary leap for our species” (Klein, 2023). Some people have built up the idea of what artificial intelligence can do on a “pyramid-scheme type of hype” (Dangerous, 2023). Much like how humans learn, AI relies on training data to recognize objects, however, “humans can do something that AI cannot. They can intuitively deduce... a previously unencountered object” (Kaust Discovery, 2023). Albeit the last frontier of AI is imagination. To study this, two categories of utopia will be examined through literature and distilled into simple prompts to assess how well AI interprets our understanding of utopian forms. These categories originate from the fundamental directions of time, futuristic and nostalgic. Simple prompts were then refined from literature on the subject and provided to the Playground 2.5 model for evaluation of compliance.

Literature

Futuristic

The following includes an analysis of the literature used to develop criteria for the playground.ai model. The literature draws from futuristic thinking utopian authors Ebenezer Howard’s *Garden Cities of Tomorrow*, Le Corbusier’s *A Contemporary City*, Paolo Soleri’s *Arcology: The City in the Image of Man*, Frank Lloyd Wright’s *Broadacre City: A New Community Plan*, Katherine McLaughlin’s article *The 6 most futuristic cities being built around the world*, La redazione di Domus article *8 floating cities between present and future*, Adam Williams article *Huge circular building runs rings around world’s tallest skyscraper*, and Reyyan Dogen’s article *The line*.

The criteria for the futuristic AI-generated image includes modular, manufactured, automated, avant-garde, trees, simple esthetic, public space, and big city. It was evaluated through several test rounds that simple, basic, and minimal commands worked the best for instructing the AI model when generating images. All the projects included in the futuristic literature wrote of some sort of automation, including public space and a simple esthetic, were avant-garde and described their utopia as a large city.

The AI model was instructed to be modular because of several futuristic thinking utopian projects including the Maldives Floating City, The Line, *Arcology*, and a project for floating houses. The Maldives Floating City project is to include replaceable parts, The Line is to be adaptable, a project for the design of floating houses describes them as “modular containers” (8, 2023), and Soleri wrote of his utopia that it would be a large three-dimensional space that was broken up into small and larger subspaces.
The AI model was instructed to make the futuristic cityscape image manufactured. This stems from the Line project which is to be manufactured and Le Corbusier in *A Contemporary City* where he wrote of mass-produced and manufactured housing to create balance, comfort, and precision in living quarters while reducing the cost of housing overall.

Trees were included as criteria for the AI-generated image to represent nature being integrated into the urban form. Le Corbusier wrote of increasing parks and open space in the city, that the entire city should be a park. Ebenezer Howard believed that nature was so important that he even included it in the title of his utopian writing, *Garden Cities of Tomorrow*. And finally, the Downtown Circle and Line projects both included a plethora of parks and nature in their site plans.

**Nostalgic**

The following includes an analysis of the literature used to develop criteria for the playground.ai model. The literature draws from nostalgic thinking utopian authors Camillo Sitte’s *The Art of Building Cities*, Charles Fourier’s *The Phalanstery*, New Urbanist principles, Karine Basset and Michele Baussant’s *Utopia, Nostalgia: Intersections*, articles on Silicon Valley’s Plan for a 55,000-acre utopia, Niklas Salmose’s *Nostalgic Utopia*, and Alan Marshall’s article *Utopian cities*.

The criteria for the nostalgic AI-generated image include communal, public space, trees, art, statues, crowd, pedestrians, and small town. All the projects included in the nostalgic literature wrote of walkability/communal space and the sense of community that they create. Public space is a principle of New Urbanism. Sitte also wrote of the importance of public space in *The Art of Building Cities* describing the town square as the hub of a town's activity. The presence of a crowd in the AI-generated image is important to show this sense of community that is formed. Pedestrians show the walkability of the town.

Trees represent the integration of nature into the town. Niklas Salmose wrote of the Paradise myth of the tree, the water, and the garden. New Urbanist principles include the integration of parks and green space within a town. The plan for Silicon Valley includes sketches of people walking along a small row of houses bordering a park with an abundance of trees.

Statues and art were stressed in the writings of Camillo Sitte to create a sense of community. He wrote of statues boarding a central community space so that they were displayed against the neutral background of buildings such as painters do in a portrait and so as not to obstruct pedestrian traffic.
Methods

A selection of software for generating an image using AI was assessed before playground.ai was selected. Some of these include Pixray, Runway, Leap Ai, getting.ai, and DeepAI. Most of them ran off Stable Diffusion, but some had a custom option such as playground.ai’s Playground 2.5 model. A series of images were run on playground.ai, including a birds-eye view, a streets-eye view, and a nightlife view. These images were then assessed on a score from zero to two. Zero meant that the criteria were not followed, one was that the criteria were followed to some extent, and two was that the criteria were mostly met. The maximum score an image could get is ten out of ten.

Criteria

Criteria were distilled from the literature and instructed to the AI model before being evaluated.

Futuristic

Parameters

Filter: None

Expand Prompt: On

Model: Playground v2.5

Number of Images Generated: 3

Prompt

A cityscape (birds-eye view/ streets-eye view/ night life), modular, manufactured, automated, avant-garde, trees, big city, simple esthetic, public space

Negative*:

- Asymmetry, Unsustainable, Blurry, Low quality, Bad anatomy, Cartoon, CGI, Anime, Grainy, Distortion, Error, Misalignment, Deformed structures, Unconnected streets

*Negative criteria are a list of parameters informing the AI model of what to exclude from the image generation process (Ahmed, 2024).
Nostalgic

Parameters
Filter: None
Expand Prompt: On
Model: Playground v2.5
Number of Images Generated: 3

Prompt
A cityscape (birds-eye view/street-eye view/ night life), communal, public space, trees, art, statues, crowd, pedestrians, small town

Negative:
- Modern, Depreciate, Uniform, Blurry, Low quality, Bad anatomy, Cartoon, CGI, Anime, Grainy, Distortion, Error, Misalignment, Deformed structures, Unconnected streets

Results

Futuristic:

<table>
<thead>
<tr>
<th>Birds-eye view</th>
<th>Streets-eye view</th>
<th>Night life</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Birds-eye view image" /></td>
<td><img src="image2.png" alt="Streets-eye view image" /></td>
<td><img src="image3.png" alt="Night life image" /></td>
</tr>
</tbody>
</table>

Evaluated Criteria:

<table>
<thead>
<tr>
<th>Criteria</th>
<th>1st Image</th>
<th>2nd Image</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public Space</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Trees</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Modular*</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Avant-garde*</td>
<td>2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Big City</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9/10</strong></td>
<td><strong>7/10</strong></td>
<td><strong>7/10</strong></td>
</tr>
</tbody>
</table>

*Modular is defined by manufactured parts of a city that could be combined in different ways.
*Avant-garde is defined as experimental, new, and innovative.

**Nostalgic:**

<table>
<thead>
<tr>
<th>Birds-eye view</th>
<th>Streets-eye view</th>
<th>Night life</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
<td><img src="image3.png" alt="Image" /></td>
</tr>
</tbody>
</table>

**Evaluated Criteria**

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Public Space</th>
<th>Streets-eye view</th>
<th>Night life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Art</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Crowd</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Small Town</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>9/10</strong></td>
<td><strong>7/10</strong></td>
<td><strong>6/10</strong></td>
</tr>
</tbody>
</table>

**Conclusion**

Some believe that AI will save us and create a utopian future, and some warn of the dangers that it poses. Much like the two fundamental directions of time, are the two beliefs surrounding AI discourse. One is that AI will be our future savior and one is that AI can only lead to a further fragmentation of our society away from a nostalgic sense of community. Nonetheless, AI is here, and it is changing the way decisions are being made. It also affects people's expectations for an ideal-utopian-world. We must plan and prepare how to ethically use AI.

From this study, it was found that long and complex prompts overwhelm and confuse AI-image generators. Rather, simple distilled prompts work best when instructing an AI-image generator. AI image generation can be a powerful tool in helping to visualize future plans for cities and create a more participatory planning process, but it can also be used to spread false information...
and sow discourse in the public. Understanding and recognizing the difference between the two is a first step in working to make AI image generation more ethical.

Work Cited


https://www.salon.com/2023/07/29/visions-how-the-quest-for-utopia-could-lead-to/


Heaven, W. D. (2023, October 5). AI is dreaming up drugs that no one has ever seen. Now we’ve got to see if they work. *MIT Technology Review.*


[https://www.gutenberg.org/files/46134/46134-h/46134-h.htm](https://www.gutenberg.org/files/46134/46134-h/46134-h.htm)


Kaust Discovery. (2023, March 31). *AI uses your imagination.* KAUST Discovery.


Le Corbusier. (1929). *A Contemporary City* [Review of *A Contemporary City*]. The City of Tomorrow and Its Planning.

[https://courses.washington.edu/gmforum/Readings/LeCorbusier.pdf](https://courses.washington.edu/gmforum/Readings/LeCorbusier.pdf)


