Process to generate Calibrating Stretched Transparency

To dig deeper into how we generated Calibrating Stretched Transparency, we explain how we collaborated with artificial intelligence (AI) and map projection tools and highlight their biases often overlooked.



Patterson Cylindrical



Peirce Quincuncial

Main stages in our creation process: (Click to drop to section)

1. Artificial Intelligence (AI) to Create the Landscapes

2. Artificial Intelligence (AI) to Face Ethical Questions

3. Artificial Intelligence (AI) to Refine the Landscapes

4. Mapping Tools Add Obfuscation and Distortions

---

1. Artificial Intelligence (AI) to Create the Landscapes

Generating the initial landscape images using neural networks (or machine-learning).

During the first stages of our project, we worked with two machine-learning tools/algorithms[1]: CLIP[2] and VQ-GAN.[3]

VQ-GAN is an "image generator"[4] using a traditional computer vision neural network. It is pre-trained on hand-labelled datasets (like ImageNet and COCO).

CLIP is a model to determine which caption is best suited for an image.

Used together, CLIP and VQ-GAN use text as an input to generate images that the machine learning algorithms predict as most closely resembling the text.

Key to our research is CLIP's capability to connect diverse natural languages and produce completely novel images that aren't within its learned data set, as opposed to traditional computer vision models that generate images close to their datasets or pre-trained networks. With this capability, CLIP is able to make the leap in developing entirely new and diverse images, which it determines are the best representations of the text given.

For example, although CLIP would have never been trained on an image described as the following, inputting the sentence “a mechanical forest in the depths of the ocean where purple plants walk on stilts” entering that description into CLIP would produce a relevant image.

Another layer of complexity added to this project is that CLIP is not trained in the traditional method of manually labeled images to create a dataset, rather it is trained on 400 million labeled images from the Internet.[5] Taking this into account forces us to consider the nature of the images and the rhizomatic effect that comes with the Internet, as opposed to manually labeled and chosen images.



Portion of the generated landscape

2. Artificial Intelligence (AI) to Face Ethical Questions

For our project, CLIP’s and VQ-GAN’s ability to produce images based on text inputs was particularly interesting for exploring the biases and modes of perception embedded in certain tools such as neural networks.

CLIP is touted as having a robustness or success rate of properly identifying images with text of around 75%[6], so we could put into question the properties and values of a relatively “accurate” neural network – especially one that is capable of producing images outside of its learned dataset. One way we explored this line of questioning was by prompting the neural network with ethical questions related to geo-engineering and AI. Words and phrases we explored with the neural networks include:

“act | someone else responsible | fails | solution | privilege | first world”
“augment | compliance | safety | utopia | trust”
“audit | diversity | solution”
“disproportion | anything that isn’t illegal goes | safety | dystopia | trust”
“act | someone else responsible | fails | solution”



Portion of the generated landscape

3. Artificial Intelligence (AI) to Refine the Landscapes

Image denoising and Inverted Image Restoration:

Further within our process, we enhanced details of our landscape images using neural networks.

Neural networks are commonly used for image restoration and generation. Denoising and increasing the resolution of images has traditionally been done through learned-prior[7] methods, where a neural network learns how an image should look through a pre-established dataset. Other methodologies like explicit prior[8], have shown the capabilities of only taking into consideration the degraded image as opposed to requiring large datasets and pertained networks[9]

Taking into account the complexities of both methodologies, we can further develop a line of questioning where we ask how the biases within these learned networks, or possible lack of bias in the explicit prior methodologies, might affect the way we use tools and come to know the world through these tools.



Portion of the generated landscape

4. Mapping Tools Add Obfuscation and Distortions

***Projecting our images with standard mapping formulas*.**

Generating a map of an area in the world requires a systematic process of flattening a 3D surface of the globe into a 2D plane[10]. The peculiarity of this process has to do with the fact that there will necessarily be distortions in any map. There is a large amount of are many possible map projections (as seen in our project). Each projection embeds decisions about which distortions are acceptable, and which are not, for the specific purpose of the map. These distortions are represented using the Tissot's indicatrices (red circles) in figure A for the Behrmann projection and in figure B for the Mercator projection.

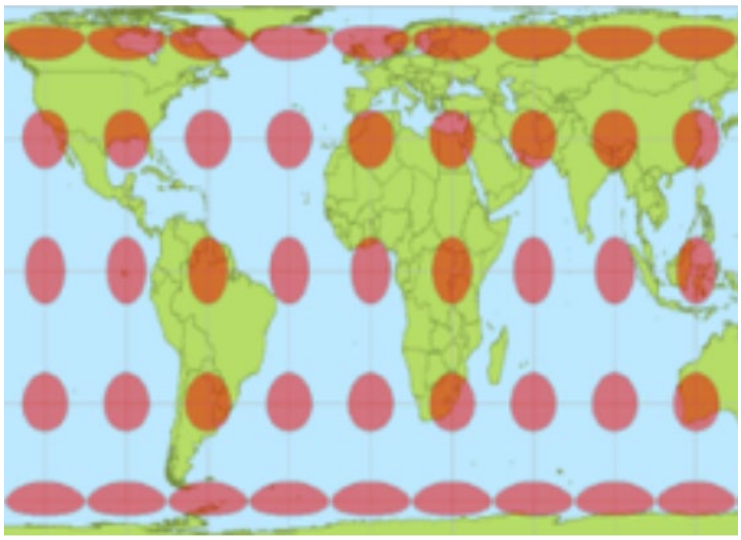


Figure A. Projection Behrmann avec indicateurs Tissot

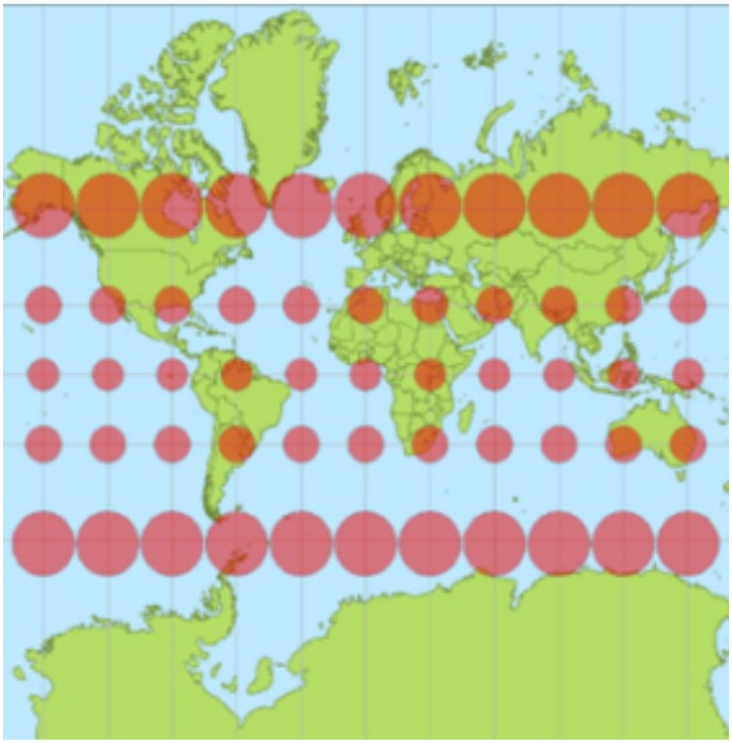


Figure B. Projection Mercator avec indicateurs Tissot indicatrices

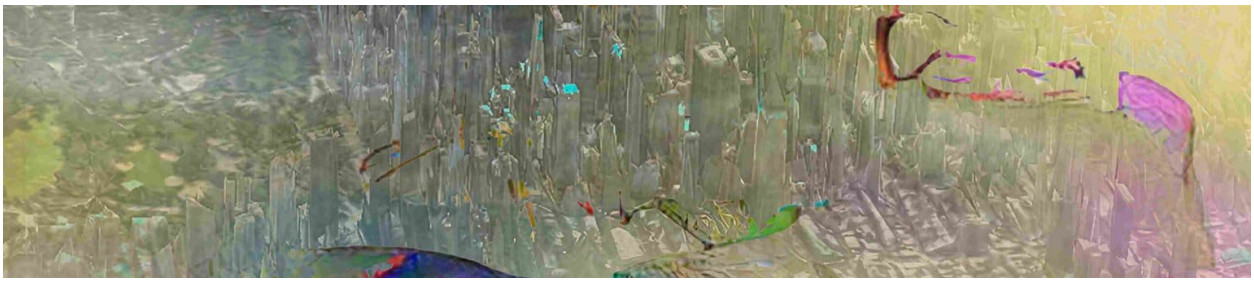
By using G.Projector, developed by **NASA,**[11] we had the capacity to project our generated landscapes through traditional projection formulas that have these chosen properties and exceptions embedded within the output of the map.



Portion of the generated landscape

Bibliography:

[1] More specifically known as neural network architectures
[2] https://openai.com/blog/clip/
[3] https://compvis.github.io/taming-transformers/
[4] https://alexasteinbruck.medium.com/vqgan-clip-how-does-it-work-210a5dca5e52
[5] https://openai.com/blog/clip
[6] https://openai.com/blog/clip
[7] https://towardsdatascience.com/demystifying-deep-image-prior-7076e777e5ba
[8] https://towardsdatascience.com/demystifying-deep-image-prior-7076e777e5ba
[9] https://towardsdatascience.com/demystifying-deep-image-prior-7076e777e5ba, https://dmitryulyanov.github.io/deep_image_prior
[10] https://en.wikipedia.org/wiki/Map_projection
[11] https://www.giss.nasa.gov/tools/gprojector/



Portion of the generated landscape

Calibrating Stretched Transparency was generously supported through the Scotiabank Fund for AI and Society at the University of Ottawa AI + Society Initiative, the Faculties of Engineering and Arts (uOttawa), and the Social Sciences and Humanities Research Council of Canada.