

ALGORITHMIC PORTRAIT

Final Report

by

Audrey Coulombe

Link to github repository:

https://github.com/AudreyCoulombe/CART451_SemesterProject

Link to live action screen recording:

https://drive.google.com/drive/folders/17MvomO5MIIE1nBnOT4AtC_2XzwmsatQ?usp=share_link

Algorithmic Portrait is an app that allows users to visualize how they are perceived by algorithms and, more importantly, how AI perpetuates and reinforces biases and stereotypes in image generation. The text to image AI DALL-E Mini was chosen for its accessibility and its use of raw data.

Accessibility: DALL-E Mini is not only free, it can also run from a [backend server](#) with Google Collab to **reduce the need in computer power**. However, in November 2022, the [Github repository](#) changed the AI used and replaced DALL-E Mini with [Stable Diffusion V2](#).

Use of raw data: DALL-E Mini was trained with unfiltered images, which is particularly relevant to my project since stereotypes are even more likely to be visible with such data.

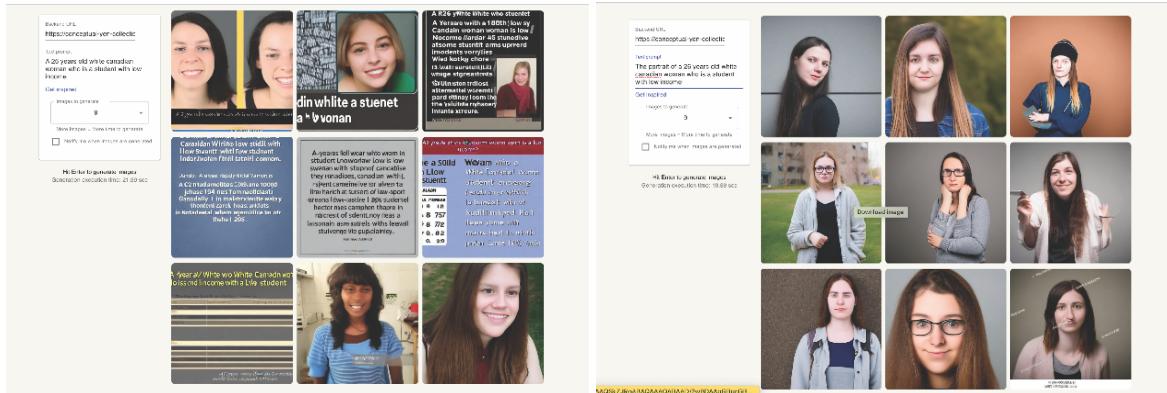
When trying out the application, users are first invited to take a quiz to “find out how algorithms see them”. At first, my goal with this quiz was to recall quizzes on social media that invite people to “Learn which Avenger you are!” (quizzes whose primary purpose is to collect personal data). However, as my project took shape, I noticed that the more information my prompt for DALL-E Mini contained, the harder it became to see the AI biases. Indeed, the most effective way to make encoded stereotypes visible is to give a certain degree of freedom to the AI with a general prompt and then observe its assumptions. Having too many parameters to take into account makes bias analysis too complex to be effective. Since this project is intended to be a tool for raising awareness, highlighting and documenting the biases embedded in AI, I decided to keep only six key characteristics that I felt were most impacted by stereotypes: (1) the skin color, (2) the gender, (3) the country of origin, (4) the income, (5) the age and (6) an activity.

When the user is done answering the six questions, the application uses the answers to create a description of the person's portrait that goes like this:

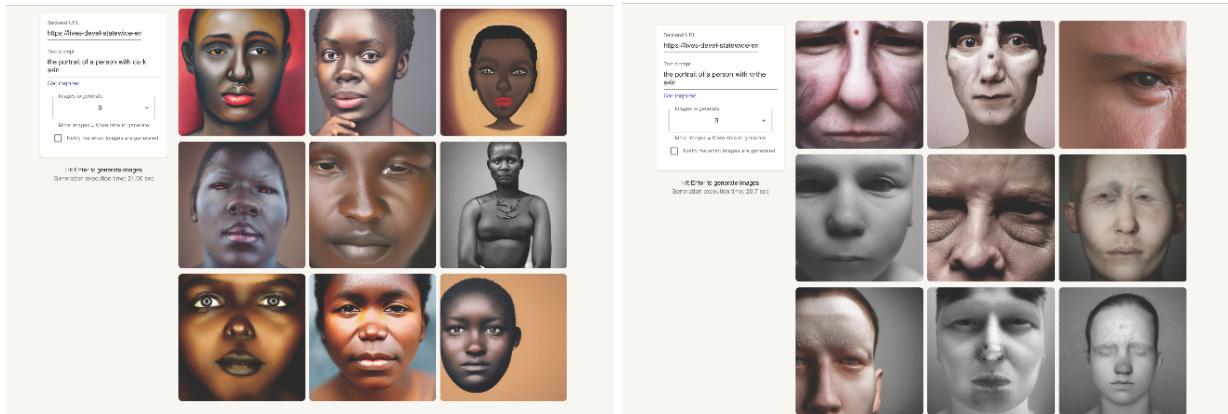
"The portrait of a " + **age** + " years old " + **color** + " " + **origin** + " " + **gender** + "
with " + **income** + " income who is " + **hobby**;

This description is then used as a prompt to generate an image with Stable Diffusion V2 backend server. Multiple tests were performed to arrive at this description structure.

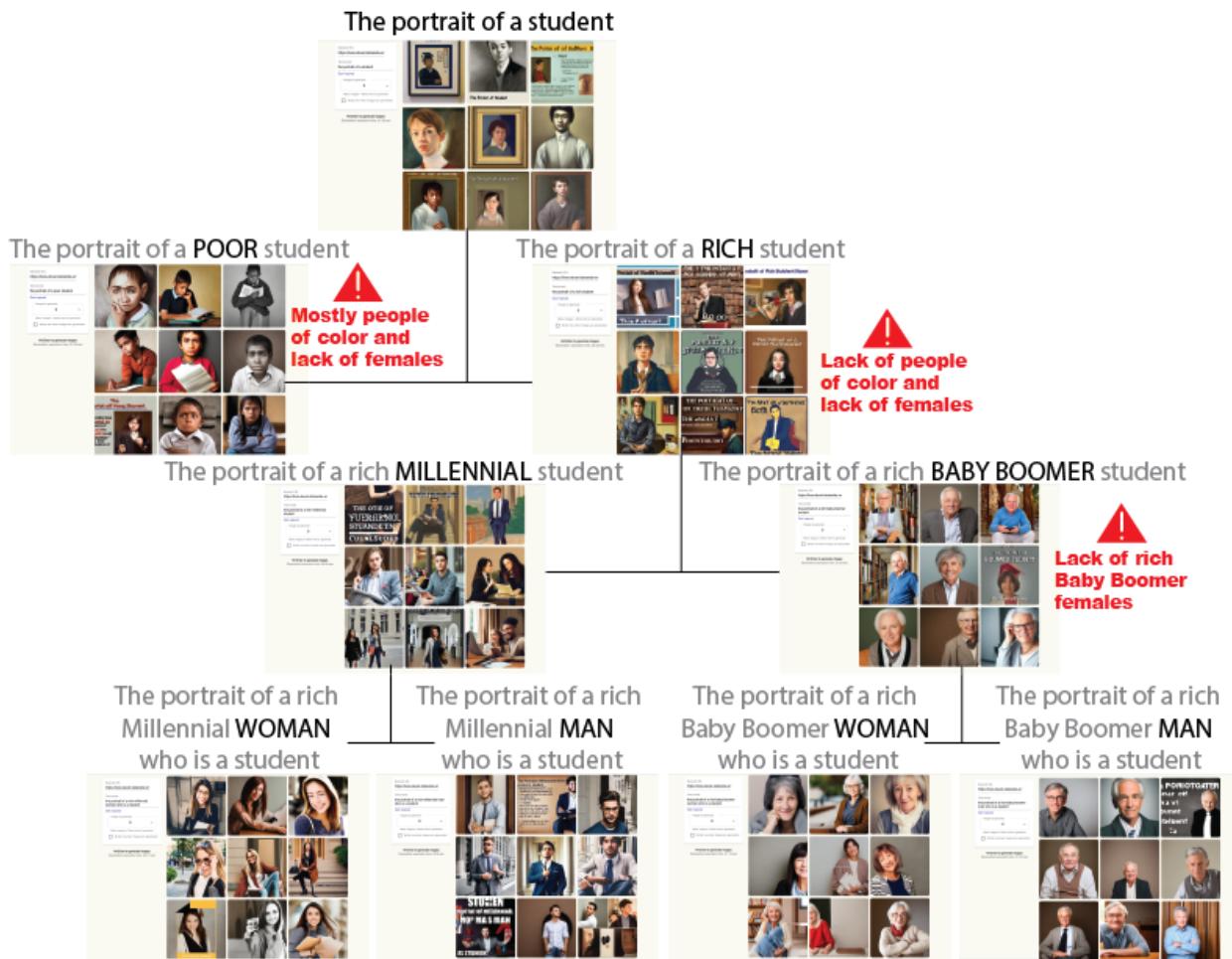
First, it is necessary to have "the portrait of" at the beginning of the description, otherwise the returned images consist mostly of text-like shapes. For example, below, the images on the left were generated from the prompt "A 26 years old white Canadian woman who is a student with low income", while the images on the right were generated from the prompt "The portrait of a 26 years old white Canadian woman who is a student with low income".



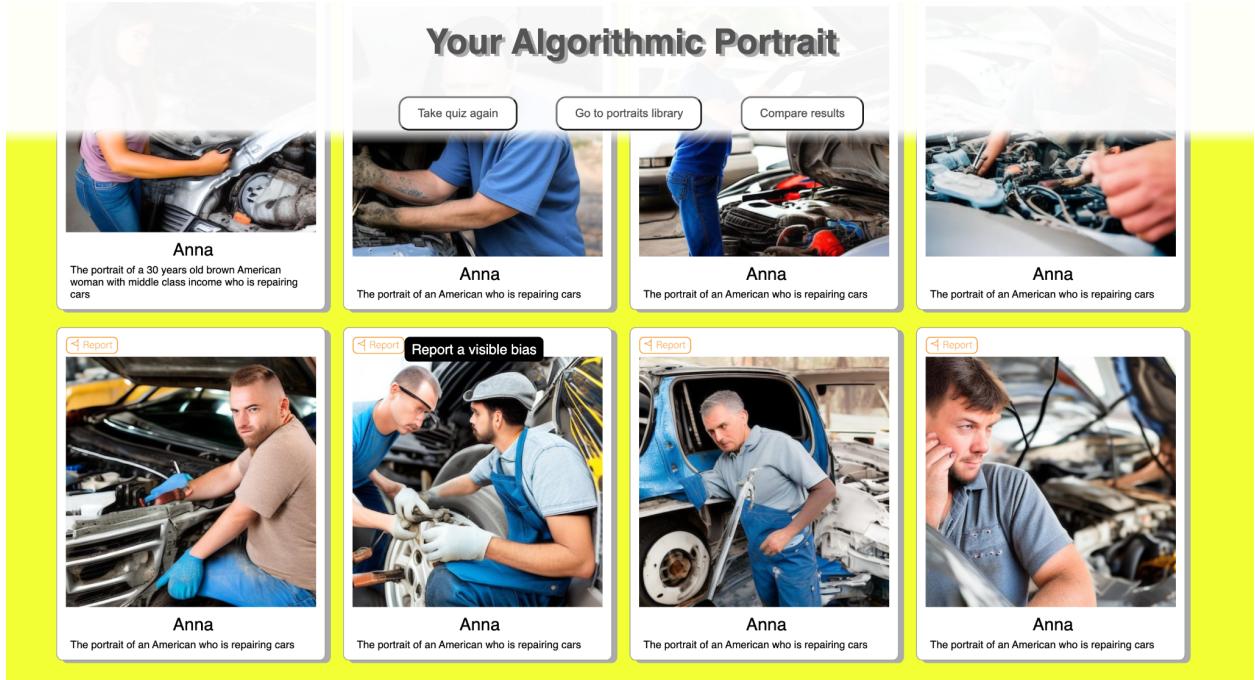
Another example of fine-tuning that was necessary with the formulation concerns skin color. Indeed, I noticed that it is best not to mention the word "skin" when referring to skin color, otherwise the AI puts too much emphasis on the skin and less on the whole portrait. Below, the images on the left were generated from the prompt "The portrait of a person with dark skin" while the images on the right were generated from the prompt "The portrait of a person with white skin".



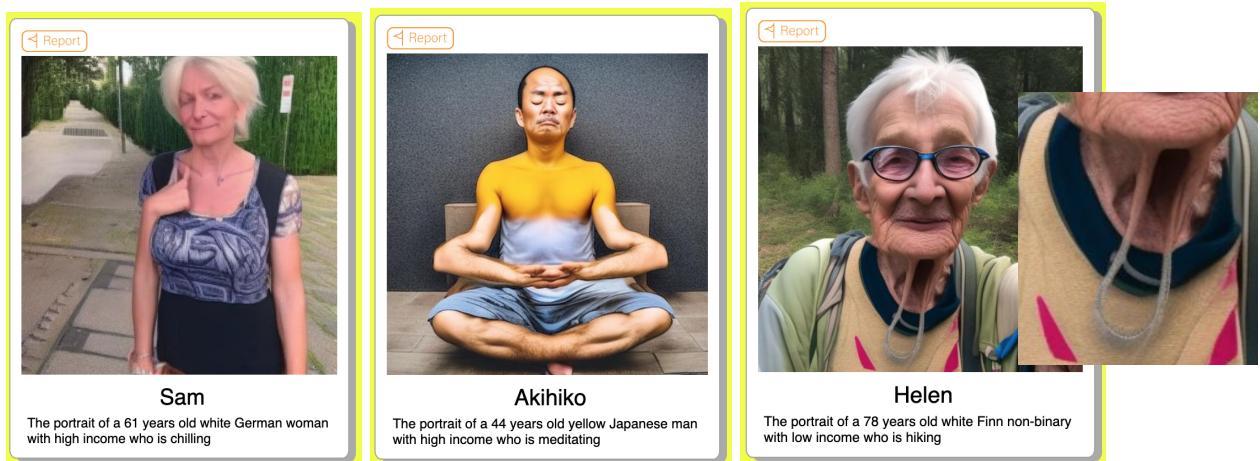
Below is an example of tests I made to help me understand how each parameter influences the outcome.



While I was experimenting with the AI, I realized that it would be easier to identify the biases if we could change or remove a descriptive parameter in the prompt and then compare the resulting images together. Indeed, this allows the user to better understand the impact of each answer/parameter on the image. I therefore added this feature to the code. In addition, this feature allows users to compare multiple results for the same prompt and to check whether a particular assumption is repeated. For example, you can see below that, when gender is not mentioned, the prompt "The portrait of an American who is repairing cars" always returns images of males.



It is also common that the generated image is completely off or uncanny. However, these images are the ones that have generated the most excitement and interest from the testers. In itself, this misinterpretation of the person's description is a good demonstration of how analyzing data without context can lead to misclassifications, and how algorithms don't see people as the humans they really are, but rather as a set of statistics and patterns.



While the project is not exactly as I envisioned it when I wrote the proposal, it is still very close and, in my opinion, even better. Indeed, I think it was a good choice to remove elements like the endless quiz and other elements reminding of the massive collection of personal data and the opacity of the algorithms, and instead to be transparent with the purpose of the app and how it works. In this way, the app feels more like a fun straight forward tool for documenting AI biases than a simulation of what we already face and hate every time we use a computer (like the opacity of used algorithms and the sneaky data collection). In addition, adding the possibility to change a parameter in the description and compare the resulting images really helps to identify bias. To polish and finalize the project, however, I would still need to add an “About” section explaining what algorithmic biases are, where they come from and what their impact is. Moreover, I would need to allow users to filter portraits based on selected bias (keywords).

Looking back, I learned a lot about how algorithms work and how to manipulate them to get the desired result. In addition, I now have a better understanding of how stereotypes can be perpetuated and reinforced by AI. Another unexpected but interesting learning is about copyright. Indeed, many images clearly reproduce watermarked image patterns, which are essentially used to protect copyrights and claim ownership. Among these images, we can clearly see the watermarks of Getty and Alamy, among others.

Screenshots and annotations

Quiz.html
The user has to answer seven questions

How do algorithms see you?
Take this quiz to find out!

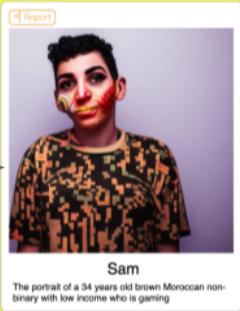
What is the color of your skin?

[Next](#)

Quiz.html
Once the user answered the seven questions, they get back their algorithmic portrait

Your Algorithmic Portrait

[Take quiz again](#) [Go to portraits library](#) [Compare results](#)

 Sam
The portrait of a 34 years old brown Moroccan non-binary with low income who is gaming.

Options to take quiz again, go see the portraits library or modify the prompt to compare results together

Your Algorithmic Portrait

Take quiz again Go to portraits library Compare results

* Tip: Try changing one feature at a time to better understand the various stereotypes associated with it. You can also remove some parameters to see the AI's assumptions.

The portrait of a 34 years old brown Moroccan non-binary person with high income who is gaming

Generate a new image

Sam
The portrait of a 34 years old brown Moroccan non-binary with low income who is gaming

Sam
The portrait of a 34 years old brown Moroccan non-binary person with low income who is gaming

The “Compare Results” option to modify the prompt and compare results together

Loading spinner animation while waiting for the AI to generate the image

CSS animation by MattIn4D
<https://codepen.io/MattIn4D/pen/DQOGBo>

Your Algorithmic Portrait

Take quiz again Go to portraits library Compare results

Sam
The portrait of a 34 years old brown Moroccan non-binary with low income who is gaming

Sam
The portrait of a 34 years old brown Moroccan non-binary person with low income who is gaming

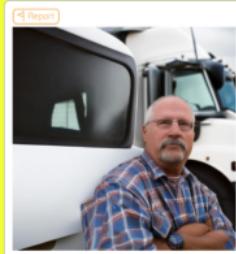
Sam
The portrait of a 34 years old brown Moroccan non-binary person with high income who is gaming

Sam
The portrait of a 34 years old brown Moroccan non-binary person with high income who is singing

portraitsLibrary.html

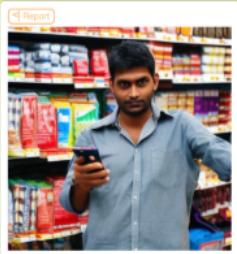
Portraits library

Age Skin color Origin Gender Hobby Income Apply Filters



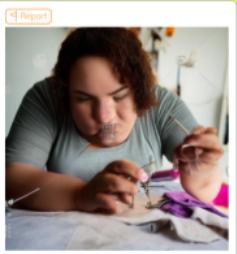
Dan

The portrait of a 58 years old white American man with middle class income who is driving trucks



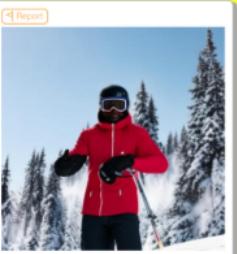
matthew

The portrait of a 32 years old brown Indian man with middle class income who is shopping



Audrey

The portrait of a 26 years old white Canadian woman with low income who is sewing



Francis

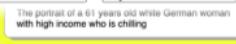
The portrait of a 41 years old black Canadian man with high income who is skiing



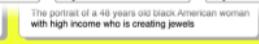
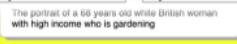
portraitsLibrary.html

Portraits library

Age Skin color Origin Gender Hobby Income Apply Filters



The portrait of a 61 years old white German woman with high income who is chilling

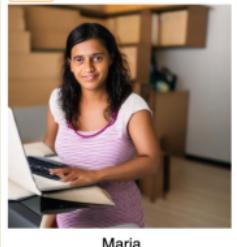


“Report a visible bias” appears when you hover the “report” button



Robin

The portrait of a 24 years old white Canadian man with low income who is playing the guitar



Maria

The portrait of a 34 years old brown Brazilian woman with middle class income who is coding



Ali

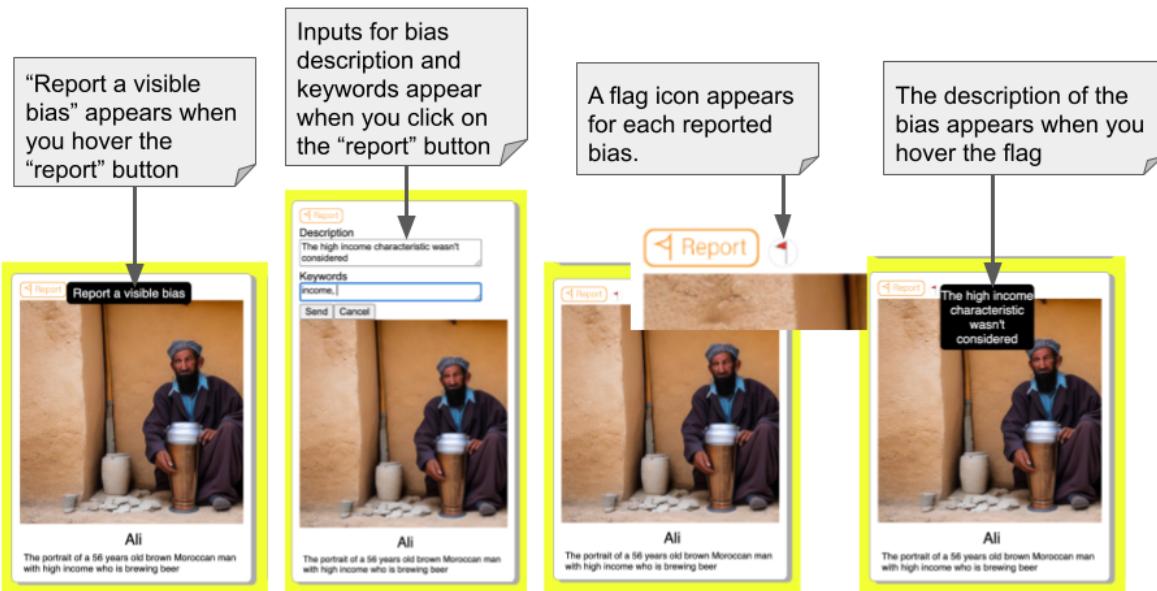
The portrait of a 56 years old brown Moroccan man with high income who is brewing beer



Thomas

The portrait of a 32 years old white Canadian man with middle class income who is gardening





The filter options

Portraits library

Age: Any, Skin color: brown, Origin: Any, Gender: Any, Hobby: Any, Income: Any, Apply Filters

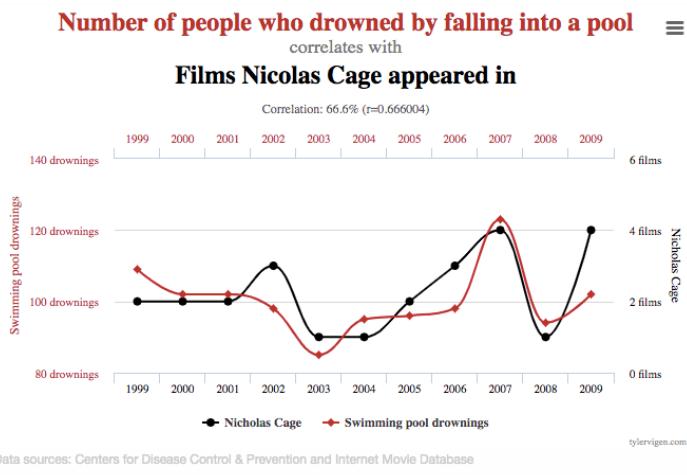
| Portrait | Name | Description |
|----------|----------|---|
| | Sam | The portrait of a 61 years old white German woman with high income who is chilling |
| | Hele | The portrait of a 68 years old woman with high income who is gardening |
| | Michelle | The portrait of a 48 years old black American woman with high income who is creating jewels |
| | Fanny | The portrait of a 20 years old white American woman with middle class income who is making bonsai trees |
| | Robin | The portrait of a 24 years old white Canadian man with low income who is playing the guitar |
| | Maria | The portrait of a 34 years old brown Brazilian woman with middle class income who is coding |
| | Ali | The portrait of a 56 years old brown Moroccan man with high income who is brewing beer |
| | Thomas | The portrait of a 32 years old white Canadian man with middle class income who is gardening |

Inspirations

Spurious Correlations by Tyler Vigen

This project directly relates to the notion of apophenia, which consists of seeing connections within random data. Indeed, *Spurious Correlations* uses algorithms to compare datasets and look for similar patterns. The correlations found are then visually communicated with charts highlighting how patterns can be seen anywhere and how misleading it can be. In fact, this experiment gave way to a lot of surprising and random resemblance. For example, a correlation was found between the number of people who drowned by falling into a pool and the number of films Nicolas Cage appeared in (see image below). Other than to amuse us and make us think about statistics differently, Vigen's erroneous correlations have little impact on our daily lives. However, it is much more disturbing to think about the algorithms that use similar methods to analyze our behavior and preferences, and thus impact our overall experience of connected interfaces by deciding what we see, who sees our content, the type of ads that target us, etc. Indeed, such algorithms are also based on pattern recognition and, as highlighted in the work of Tyler Vigen, these can lead to misinterpretation.

I found this work particularly inspiring for my project because it communicates a concerning problem behind data analysis. It is also relevant to mention that I first explored this project a few years ago, and the fact that I still refer to it today shows how much it has impacted me. Indeed, I've seen many other projects over the past few months that I don't even remember, but I've never forgotten this one. With my project, I would like to attain a similar balance between entertaining and informing people about the biases behind algorithmic analysis so that they remember that there is no such thing as neutral data and that their analysis must be taken with a grain of salt.



Source: <https://tylervigen.com/spurious-correlations>

[Classification.01](#) (2017) by Mimi Onuoha

This artwork is about the algorithmic categorisation of people. It critiques the way our society blindly perceives statistics as undeniable facts, without even questioning the processes involved. Classification.01 takes the form of neon tubes in the shape of curly brackets, placed on a wall. When two people stand in front of them, cameras combined with various algorithmic processes are used to analyze their appearance and determine whether or not they are similar. The neon tubes light up when they are classified as "similar." As often, the process behind these algorithmic classifications is opaque: viewers have no idea why they were categorized in this way or what data it was based on. Similarly, I want to reflect the opacity of the analysis made by the AI DALL-E 2. People will see certain details in their portrait and not understand where they come from. Like Classification.01, I also want to raise questions about why a certain categorization was made, as well as get people to think about how algorithmic analyses amplify the cognitive biases of humanity.



Source: <https://mimionuoha.com/classification01>

Movie [Coded Bias](#), directed by Shalini Kantayya and based on Joy Buolamwini's discovery

This film exposes the inability of facial recognition algorithms to detect dark-skinned people and accurately classify women. It highlights how the data that feeds these algorithms is as biased as humans. This is because the most privileged people, such as white men, produce the most data and are therefore overrepresented in that data. When algorithms are fed with this biased data, it results in biased algorithms that discriminate against underrepresented groups, which are often people of color and women. The film also shows how these discriminations can contribute to the amplification of inequalities and stereotypes, such as when police use biased algorithms to identify potentially dangerous people, which too often results in racial profiling. The denunciations made in this film are the basis of my criticism of algorithmic biases. Indeed, my wish is that the portraits resulting from my application shock and show clear traces of these biases.



Source: <https://www.codedbias.com/>