

The background features a series of wavy, parallel lines in shades of blue and cyan, creating a sense of depth and movement. A single, bright blue line glows and curves across the center of the image, passing behind the text.

AI Vision Librarian

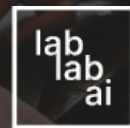
Who am i?

Isayah Culbertson

Software Engineer with a passion for LLM related technology and how it can be used to solve real world problems

github: [isayahc \(Isayah Culbertson\) \(github.com\)](https://github.com/isayahc)

linkedin: [\(1\) Isayah Culbertson | LinkedIn](#)



Problem Statement

- There is a large body of visual information that has been classified using human labor
- It is computationally expensive to perform operations on images such as clustering, and sorting

The solution

- Using Gemini multimodal functionality i input images and output a set of tags that can be used to search for the image
- the tags are vectorized and can searched for via llms
- using prompt engineering and other methods tags can be very specific and unique based on a given knowledge base.

- This technology can be used by organizations who have a large amount of unorganized visual information

Future Prospects:

- In the future as the cost of running LLMs become cheaper i would want to take videos as input. As videos have the same problem. In addition, there is a lot of insight to be gained from categorizing videos

Thanks!

