“StableSnap” - Project Proposal

Team

- Paula-Elena Gheorghe, 11926306 (e11926306@student.tuwien.ac.at)
- Daniel Martin Puehringer, 01556470 (e01556470@student.tuwien.ac.at)
- Dominik Vallazza, 11713121 (e11713121@student.tuwien.ac.at)
- Raoul Christian Wograndl, 11911109 (e11911109@student.tuwien.ac.at)

Project Idea/Scope

Motivation

When users want to capture the process of an ongoing construction site, they usually end up with huge amounts of pictures, which all show the same construction site over time. Usually, they want to select a subset of these images. If this is done manually, it takes several hours to go through all images and select the best ones. Therefore, a solution empowered by AI would be fitting.

Core idea/Minimum Viable Product (MVP)

We strive to build a platform where users can upload a large amount (e.g. 1000+ or even 10.000+) of images. With the help of AI (i.e. Amazon Bedrock) the users can request what kind of subset of pictures they would like to see and the platform will select a suitable subset of images based on the wishes of the customer. Concretely, we would like to use images of construction zones which could be useful for advertisements for construction firms. The images could be used for timelapses, in order to have an overview of the construction process. The users could say, for instance, if they would prefer images depicting sunny weather, or with multiple people in them. In addition to the described scenario, we will focus on cost efficiency and security/availability of the product.

Optional features

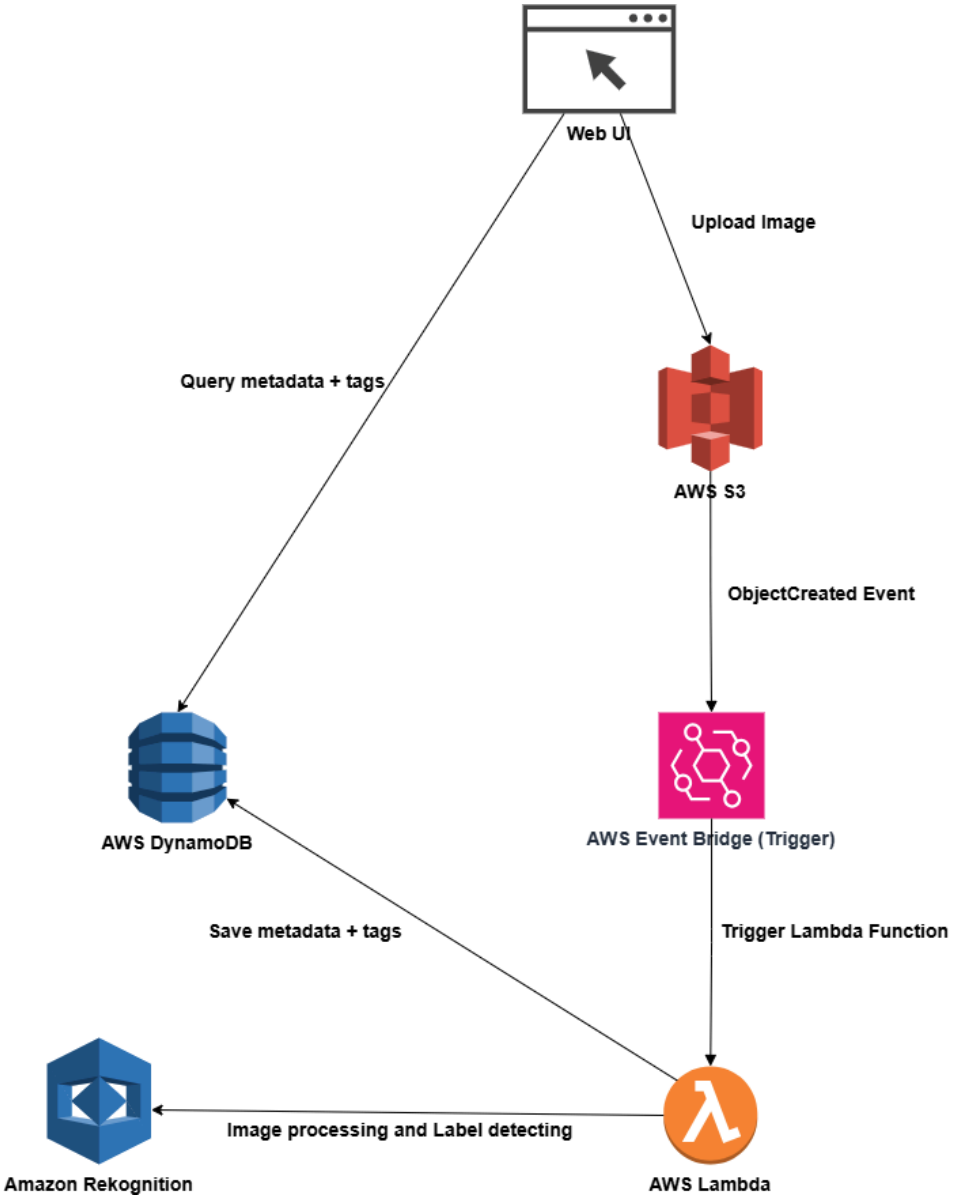
Given the high level of uncertainty, we cannot estimate the workload needed to implement the MVP. Depending on how much time is left after finishing the MVP and the recommendations from the AWS team, we want to add optional features to enhance our software. We do not intend to build all of them, only the most useful/most recommended.

- Build a second version with different (i.e. cheaper/more efficient) architecture and test which one performs better
- Build a feature which helps the user to select images which are very different from each other - e.g. images with few construction workers are usually followed by images with many construction workers to maximize the difference from each selected image

- Build a more sophisticated Frontend (with React) instead of Streamlit
- Build user management and authentication (i.e. Amazon Cognito)
- Build an integration for IoT Devices to send images to our system (directly to S3 instead of manual uploads)
- Investigate how to apply fine-tuning, Retrieval Augmented Generation (RAG) and prompt engineering
- Investigate and mitigate potential security risks (i.e. input validation,...)
- Build the application with Terraform instead of AWS CDK

Technologies

The following list of used technologies/services is not final and may be subject to change: Amazon EC2 / Amazon Lightsail / AWS Lambda, Amazon S3, Amazon DynamoDB, Amazon Rekognition / Amazon Bedrock, Amazon Cognito, Streamlit / React, AWS CDK / Terraform



Potential Risks

- Costs required for image processing
- Processing time of images
- Scope being too large for this project
- Extra time needed to learn about technology stack

Notes

Approval was received for using a received dataset with images of construction zones for training the ML model. The dataset was received from an employee of https://www.timelapserobot.com/.