Chapter 5

Atlas

In order to accomplish the crowdsourcing of accessibility scores and image annotations discussed in Chapter 4, a web based annotation tool, named Atlas, was developed as an instrument for crowdworkers to rate sidewalk images and provide additional annotations that cannot be captured by YOLOv5. This chapter discusses the annotation process of the crowdworkers as they use Atlas along with its system architecture.

5.1 System Architecture

Figure 5.1 shows the system architecture of the study. The data collection and processing scripts are separated from the main system but all data collected will be stored in a single database that Atlas will also connect to.

In terms of storage, we use Azure Storage Explorer to store all raw images and blob data. MongoDB is used to store all data collected from Atlas, such as sidewalk accessibility scores, annotations, and participant information (non-personally identifiable information) MongoDB is also used to store the links to the pre-labeled images and their annotations from YOLOv5 and OCR (ResNet-101).

For the backend of the Atlas system, we use Next.js API routes for the API responsible for querying the data and saving the user inputs to the database. These API routes retrieve the pre-labeled images from MongoDB, authenticate users, register users, and save user annotation to the database.

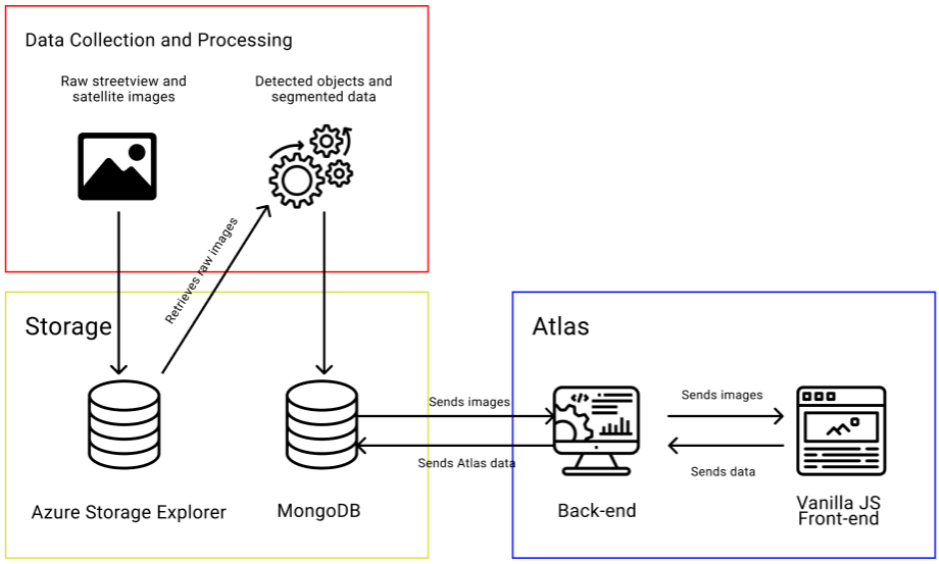

Figure 5.1: Diagram of Atlas System Architecture

For the frontend, we use Next.js and Tailwind CSS for our styling. Next.js is an open-source development framework built on top of Node.js enabling React based web applications functionalities such as server-side rendering and generating static websites. Tailwind CSS is a utility-first CSS framework that excels in rapid styling of HTML elements. We built the annotation tool by building on top of React Picture Annotation and customizing it to our needs.

MongoDB was used as the database to store the data needed and obtained from Atlas. The database contains three tables, namely *User*,*Images*, and *Annotations*. The User table contains all relevant information in building a simple demographic of our crowdworkers and user credentials for authentication. The images table contains data on pre-processed images: this include the blob link url, longitude and latitude, direction, city, and the bounding boxes for the objects that our computer vision model detected. Lastly, the Annotations tables are where user annotations are stored when they contribute to our platform. Annotations are connected to the User table to connect an annotation to a specific user. In addition, an annotation also has two relationships to the Image table: here we get the imageID and the detected objects the user selects. The exact attributes of each document are presented in Figure 5.2.

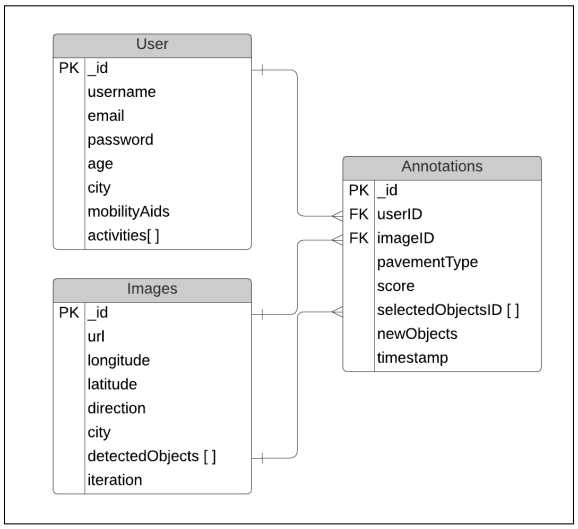

Figure 5.2: The Entity Relationship Diagram of the database schema used to store image and annotation data in Atlas

5.2 Annotation Process

The crowdsourcing platform, Atlas, is a tool where volunteers could score sidewalk accessibility, identify sidewalk objects that contribute to the score that they gave, and add new labels on objects that were not included on the pre-labeled images. This tool is based on a similar project called Project Sidewalk, wherein volunteers can validate curb ramp levels, rate them, and give comments and provide categorical tags on them.

5.3 Registration

For Atlas, users can access our online platform where they can register an account with their email. When creating an account, the user will be asked for the following participant information: *city of residence*, *age*, *use of mobility aid (wheelchairs, walkers, canes)*, and *frequency of using sidewalks in a week*. We collect this information in order to know the demographic of our users. Additionally, we want to explore the relationship of user demographic and the scores they provide. The account will be used by Atlas as a tracker on the images that the volunteer annotated. Additionally, the email collected from the users is strictly used for means of communication.

5.4 Identifying Objects

After creating an account, the volunteer can proceed to score images on Atlas. Annotators can select among 5, 10, 15, and 20 as the number of images they choose to annotate. This is to provide them with a sense of commitment and completion after they finish their annotation session.

Next, they will be shown an image containing objects on sidewalks with bounding boxes and tags annotated on them. From this point, the volunteer can do two things. They may click on the pre-labeled objects and identify if they are obstructions or not, and they may also choose to annotate more objects that they think contribute to the accessibility of the sidewalk by adding bounding boxes on objects that were not included in the pre-labeled images. They are then asked to select the class or label for these bounding boxes from a list of objects (refer to Section 4.2.1). If the object is not present in our list of objects, it means that the object is easily movable by a human being (e.g. tables) so we don't consider it as an obstruction. However, an *Others* option was also provided to catch potential items that may be considered by other as an obstruction.

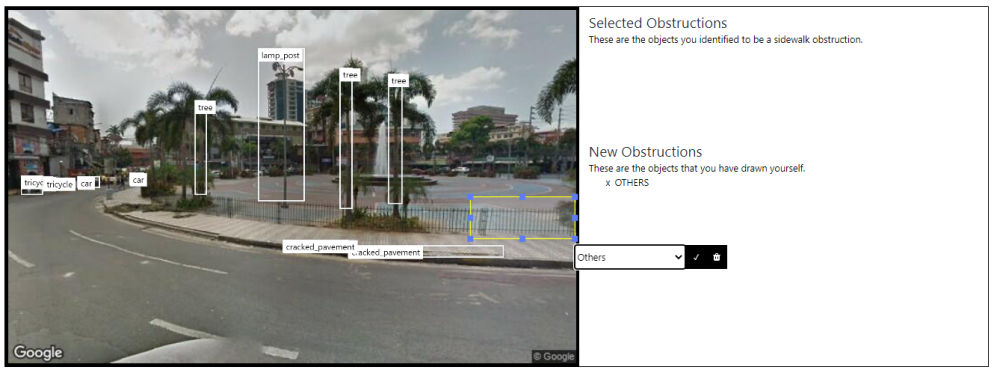

Figure 5.3: A screenshot of the Atlas prototype showing how to add a new annotation and label from a predetermined list of objects

5.5 Scoring Sidewalk Accessibility

Once participants finish annotating objects, they will then give the image a sidewalk accessibility score. Participants may select a score ranging from 1 to 10 based on their own perception of accessibility as seen in the given image, and end the annotation process by selecting the surface type of the sidewalk in the image as shown in Figure 5.4. To prevent any biased scores, did not provide any

preliminary information or expectations of what a given score should like based on the street view image. We utilize a 10-point Likert scale as a means of capturing the widely varying levels of accessibility in urban streets. Similarly, Naik, Philipoom, Raskar, and Hidalgo (2014) measures the perceived safety of sidewalks in their work, Streetscore, using a 10-point scale. The reason being that the street level images in the dataset, when compared to each other, contain different elements which improve or reduce the perceived safety of the location in the image. These elements provide a more precise attribution to the indicated Streetscore when measured on a 10-point scale. Additionally, for this study, we are focusing more on the pedestrians' perception of accessibility and trying to figure out what contributes to an accessible sidewalk. We would like to note that the participants only rate the accessibility of the sidewalk in the image presented to them. We did not consider sidewalk continuity. Additionally, images that contain sidewalks on both sides of the road are given a single score; due to the fact that sidewalks that are near each other usually have a consistent design.

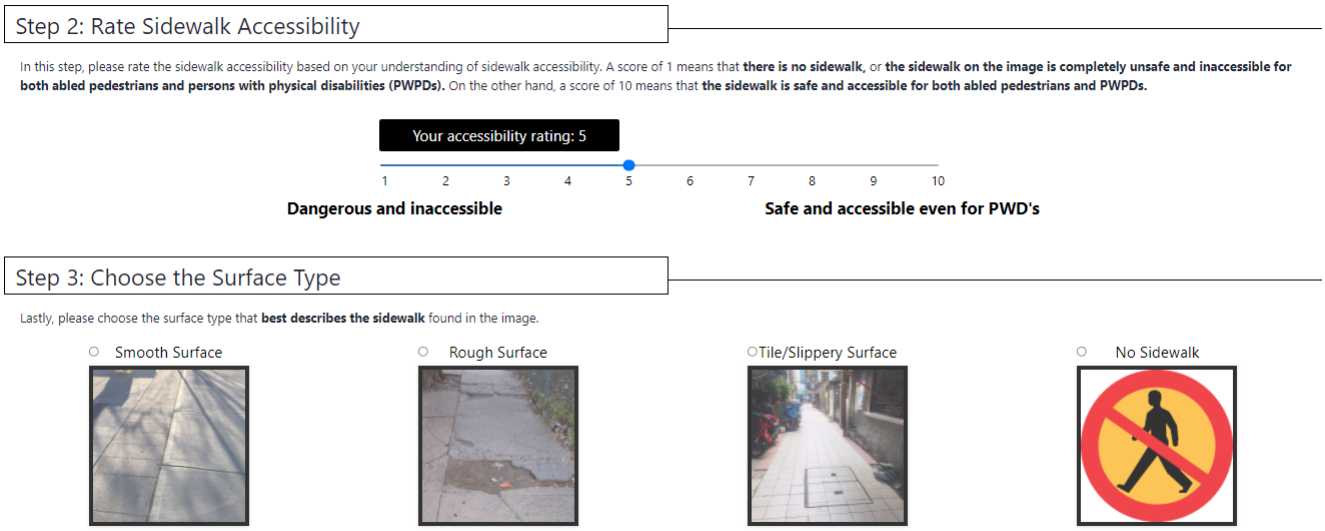

Figure 5.4: A screenshot of the Atlas platform showing the accessibility score slider as well as selecting the surface type

After giving a sidewalk accessibility score, they can proceed to the next sidewalk image and repeat the scoring and annotation process. The annotation process is designed to be simple and convenient for crowdworkers to encourage them to score as many images as possible. The simple design also contributes to the speed of data collection so crowdworkers will not feel overworked from annotating multiple images.