**JAD #1 Meeting Agenda**

**When & Where:** February 5th, 2020 ~ 10:00am - 11:00am @ SBSB 3133

**Attendees**: Dr. Wesley Schultz, Dr. Shaunn-inn Wu, Stoic Solutions Team

**10:00am - Formal Introductions**

**10:05am - Brief summary of goals for Dr. Schultz’s proposed system**

Goal #1: Collect data (360 Google Street Images)

Goal #2: Process those images through ML system

Goal #3: Collate the output data (includes Geographic Area & Standardized Level of Litter)

Goal #4: Make the data accessible to the UI team

**10:10am - Main Agenda Items and Questions**

**Jordan**: Are you aware of any existing infrastructure from the last group (AWS Server, Trained ML algorithm, Image dataset)? AWS pricing/costs? Is it funded?

**Connor**: Must/should/want to have? What would you definitely not want to see?

**Brian:** The previous group proposed standardizing the level of litter. Do you have any preferences for this standardization? (e.g. This ties into how we want to quantify the litter by a given region). Given the large amount of images that will need to be processed, is it required that we **do not** store the images on the server?

**Matthew**: Can the general idea of a “360 image” be broken down into 4 photos: front, back, left, and right? If this is the case, do we think the front and back images may not be necessary since they will only depict the forward and rear view of the street? This would cut our network usage. What is the time to process an image?

**Adam**:Can we start with the city of San Marcos at the start then, work our way outward if we refine the algorithm to a point where it is utilizing Google Streets and gathering data without us providing images to it?

**Adam**:Since we will be setting up the algorithm again and running it on our own, do you have an idea of how long it took the algorithm to start catching on to picking out trash in pictures?

**Jerry:** In regards to the output data that we generate after running the 360 google street images through the ML system: We will need to create a database that will be accessible to the UI team. We can host this database on our AWS server. How soon will the UI team need access to this database? Or is it the responsibility of the UI team to host and design a database? In this case, we would only be responsible for updating it with the data we generate based on their criteria. Should we meet with the UI team directly?

**Chris**: Any immediate UI end requirements before we meet with the UI team?

**Ricky:**Are we ok with updating the database annually, given that google updates images on a yearly basis?

**10:50am - Summary**

**11:00am - End meeting. Thank Dr. Wesley Schultz and Dr. Wu for their time.**

**JAD #1 Meeting Agenda**

**When & Where:** February 5th, 2020 ~ 10:00am - 11:00am @ SBSB 3133

**Attendees**: Dr. Wesley Schultz, Dr. Shaunn-inn Wu, Stoic Solutions Team

**10:00am - Formal Introductions**

**10:05am - Brief summary of goals for Dr. Schultz’s proposed system**

**10:10am - Main Agenda Items: Q/A**

* Existing infrastructure and pricing.
* Priorities/Wants and Needs.
* Preferences on quantifying litter in a given region.
* 360-Image breakdown.
* Start small then go further.
* Previous algorithm time-frame.
* Image storage.
* Overall functionality, time-target, and UI team integration.
* UI end requirements.
* Update frequency.

**10:50am - Summary**

**11:00am - Meeting Adjourned!**

**THIS PAGE IS STRICTLY** to brainstorm questions/matters we’d like to ask Dr. Schultz:

JAD#1 Question Brainstorming (Delete after final list is decided)

* Ask Dr. Schultz for confirmation of previous groups Github account. [Litter Detection](https://github.com/isaychris/litter-detection-tensorflow)
  + Previous group members: Chris B., Kathy M., Greg S., Malia T., Andrew Z., Vanessa J. from 2018 CIS 490 class.
* Since the project is looking for a fairly high degree of granularity can we agree to start with just the city of San Marcos? If we can get it working for a single city then it shouldn’t be much more work to expand it to the county level. (This would have to be confirmed with KAB UI group so they know that the dataset will only exist for San Marcos, but that they should design a modular UI so that data can be aggregated for each city.)
* There are many metrics that could be derived from the ML algorithm and data that is available for each city. As an example: **(If anyone can think of other metrics please add them here, also please double check my logic on calculations.)**
  + City of San Marcos Area: 24.36 sq mi (128,620.80 sq ft)
  + CSM 2018 Population Estimate: 96,847
  + (Hypothetical) ML Algorithm Pieces of Litter: 21,000
  + (Hypothetical) ML Algorithm Images Processed for CSM: 10,000
    - (Metric #1) **Litter per sq ft**: 21,000 / 128,620.80 = .1633 pieces of litter per sq ft.
    - (Metric #2) **Litter per person**: 21,000 / 96,847 = .2168 pieces of litter per person.
    - (Metric #3) **Litter per image**: 21,000 / 10,000 = 2.1 pieces of litter per image.
    - (Metric #4) **Litter Saturation**: Litter per sq ft ( .1633) \* 100 = 16.33%
  + Are any of these metrics especially appealing to Dr. Schultz? (Note: Since street view images only account for the area where roads exist it’s not a perfect measurement when using a city’s total area. Also, population and landmass would either have to be hard-coded into a table entry for that city or somehow programmatically pulled from another data source.)
* Will we be running this system to be entirely on AWS? **(Jerry)**
  + How shall we go about setting it all up on AWS?
  + Any constraints/Issues with using AWS(e.g. Slow server)?
  + Technical constraints/pricing constraints (**Matthew**)
* My Current understanding of the *general* steps we will need to take **(Jerry)**

1. Collect data (Google Street Images) with the use of google API
2. Format this data such that it will work well when fed into ML System
3. Capture the output (Numbers that categorize the level of litter)
4. Push all results to a designed database (that will in turn be used by the UI team)

* How frequently do we want the data to be updated ? **(Jerry)**

**-maybe pull annually**

* Give our idea of the project and ask if that represents what he was expecting.
* Now that we have a better understanding of the project, what is your exact vision for this project (**Enrique**)
  + (I know he wants a quantitative number) But from my understanding, the program runs a picture through rcnn, and once the objects are found and marked as litter, there is a quantitative number that is returned 1-4, (4 being more than 9 objects of trash in the picture) but what quantitative number does he want? Because their algorithm already counts how many objects are detected as trash, and just rates the amount of trash in the picture.
* What are your must haves, and wants for this project. (**Enrique)**
* We would need a database to store the quantitative number, would the school give us access to theirs? Or how would we go about acquiring one. (**Enrique)**
* How did the previous team get their access to the aws server? The ubuntu server they used previously is not free.

--Should we run a batch of images through the ML system, extract the needed data (standardized level of litter, and location, etc.) and then delete the image files? This will save a great deal of memory, but is it possible?

Brian and Jordan will use this area during the meeting to fill in details discussed regarding the project. (Don’t print this page for Dr. Wu and Dr. Schultz)

# Documentation From JAD#1 (Brian)

# Documentation From JAD#1 (Jordan)