



Application No: 2024/0334

Submitted on: June 8, 2024

A Novel Approach To Urban Smellscape Mapping For Decision Support
System

Submitted by

Rakesh Kumar Donthi
Assistant Professor
HYD-GST-CSE

Principal Investigator

Rakesh Kumar Donthi

Assistant Professor

HYD-GST-CSE

Email: rdonthi@gitam.edu

Mobile: 7816053329

DOB: 1985-02-21

Employee ID: 600818

Co-Investigator

Image	Name	Designation	College	Email	Mobile	EMP ID
	Dr. Mahender Reddy Chilukala	Assistant Professo r	HYD - GST - CSE	mchiluka@git am.edu	9908980700	600810

Proposed Research work

Title of the Project :

A Novel Approach To Urban Smellscape Mapping For Decision Support System

Broad area :

Smell significantly influences our perception of places, yet it has been largely neglected by urban planners and scientists due to the challenges in recording and analyzing it at scale. Urban odors were often linked to factories, tanneries, open sewers, and inadequate waste disposal systems.

Project Summary :

Modern cities are characterized by a complex mixture of smells that include not only the traditional industrial and waste-related odors but also a variety of other sources such as food, traffic, green spaces, and cultural activities. This diversification is partly due to the increase in urban population density, the rise of street food culture, and the proliferation of urban green spaces.

This study investigates the olfactory composition in urban environment (eg. GHMC), focusing on the dominance of distinct smells and their far-reaching effects on the city's green spaces and overall urban well-being.

This study explores using social media data to map city smells, collecting georeferenced picture tags from Flickr and Instagram and tweets from Twitter(X) of urban area (GHMC).

This study captures the urban smellscape from social media data to map city smells and finds correlations between specific smell categories and governmental air quality indicators, validating the approach.

Objectives :

- Collect smell-related words and create the first dictionary for urban smell.
 - Collect Georeferenced picture tags from Flickr and Instagram, and tweets from Twitter. Use social media data to map city smells reliably.
 - Field Survey on different odor smells in the city
 - Identify odor smell places in the urban area and capturing the entire urban smellscape on the map
 - This helps for better urban planning, proactive problem-solving and improves public health
-

1. I have done my Post Doc in the following Irish Research Council project

Green Space Engage

Led by Associate Professor Michael Lennon (School of Architecture and Environmental Policy) with Professor Michela Bertolotto, Dr Gavin McArdle and Dr Aura Istrate

The 'new normal' of remote working presents an opportunity for enhanced urban health and wellbeing by localizing work, services and recreation in walkable neighborhoods. However, this pandemic-transformed context also poses a challenge for urban planning and design as the increased demand for public green spaces during the pandemic outlasts the crisis. This project directly tackles this problem by harnessing and integrating a range of social science and computer science research approaches to produce an innovative knowledge cocreation methodology. Grounded in a novel theoretical approach, a unique coalescence of research methods will be mobilized with local communities and key decision makers to co-create new knowledge on green space perceptions. This integrative approach will open new lines of research at the interface of public health, planning, geography and landscape architecture. This new knowledge will also inform the co-production of planning, design and management guidance for public urban public urban green spaces in response to evolving patterns of living that have intensified and diversified the demands placed on such spaces as local anchors for maintaining and enhancing health and wellbeing.

2. The following paper published on related work

Covid-19 Vaccination Classification of Opinion Mining with Semantic Knowledge-based Decision Making.

Nikhila Polkampally¹, D. Rakesh Kumar², G. Soma Sekhar². International Journal of Web & Semantic Technology (IJWesT) Vol.13, No.3, July 2022, DOI: 10.5121/ijwest.2022.13301 1.

Methodology

This study explores using social media data to map city smells, collecting georeferenced picture tags from Flickr and Instagram and tweets from Twitter(X) of urban area.

Opinions are extracted from social media is processed using natural language processing.

This helps to identify

- Different types of odor smells identified by public
- sentiment (positive, negative or neutral) of each sentence
- frequency of a word mostly used
- Word clouds to get quick idea about overall opinions
- Most used Combinational words

This is implemented in the following procedure

1. Two platforms are chosen from social media twitter and google maps.

- a) Twitter data: it is collected from twitter through academic research login access where it allows analyse complete twitter data as per the requirements
- b) Google maps data: the reviews from google maps are collected manually from each park in google maps.

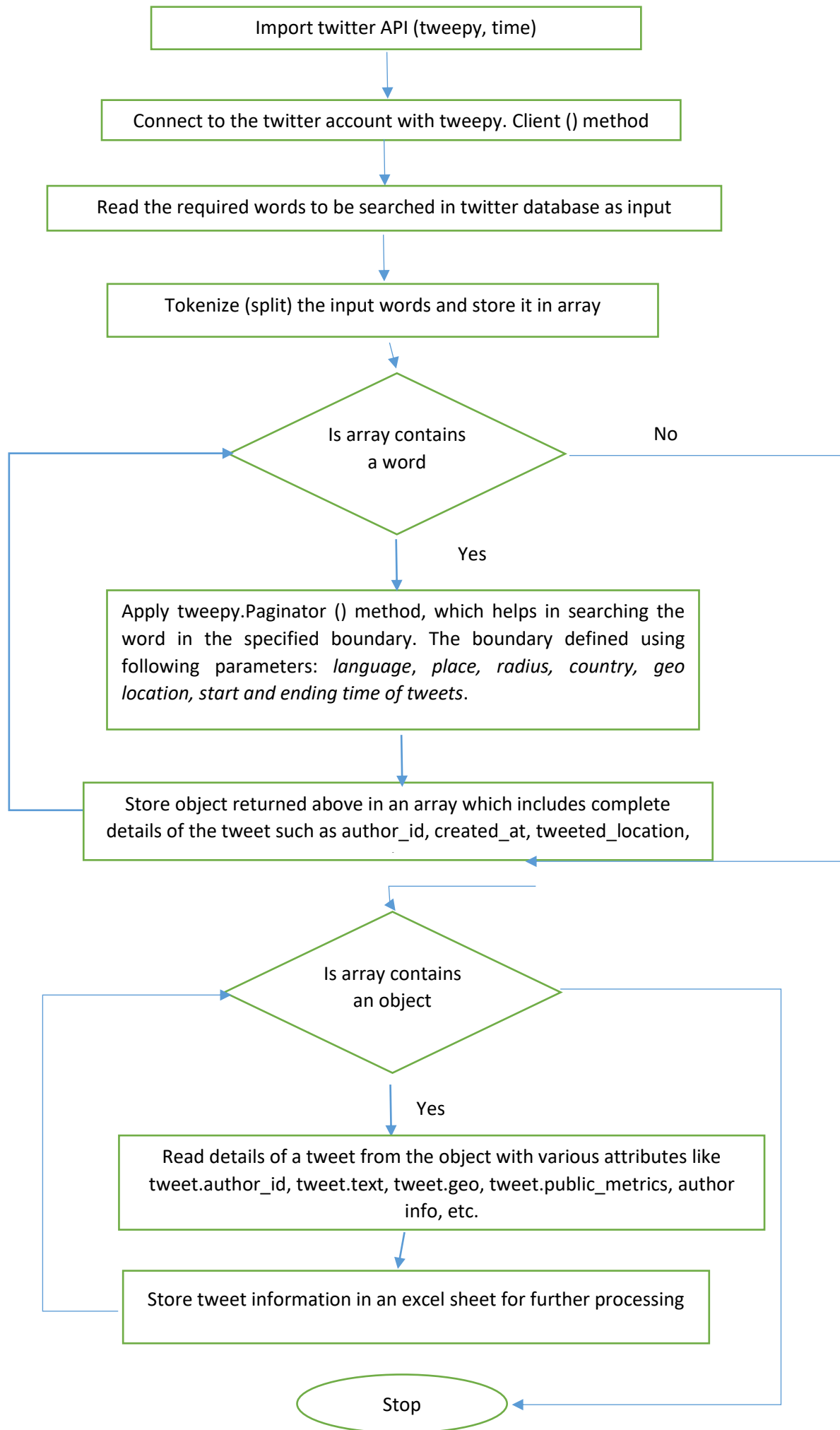
2 a). Data extraction from twitter:

Data is extracted using python code, where it uses twitter API to extract tweets from specified regions and store in excel sheets. (annexure I).it is debugged on anaconda 3.0 IDE.

Following library packages are used in extracting data:

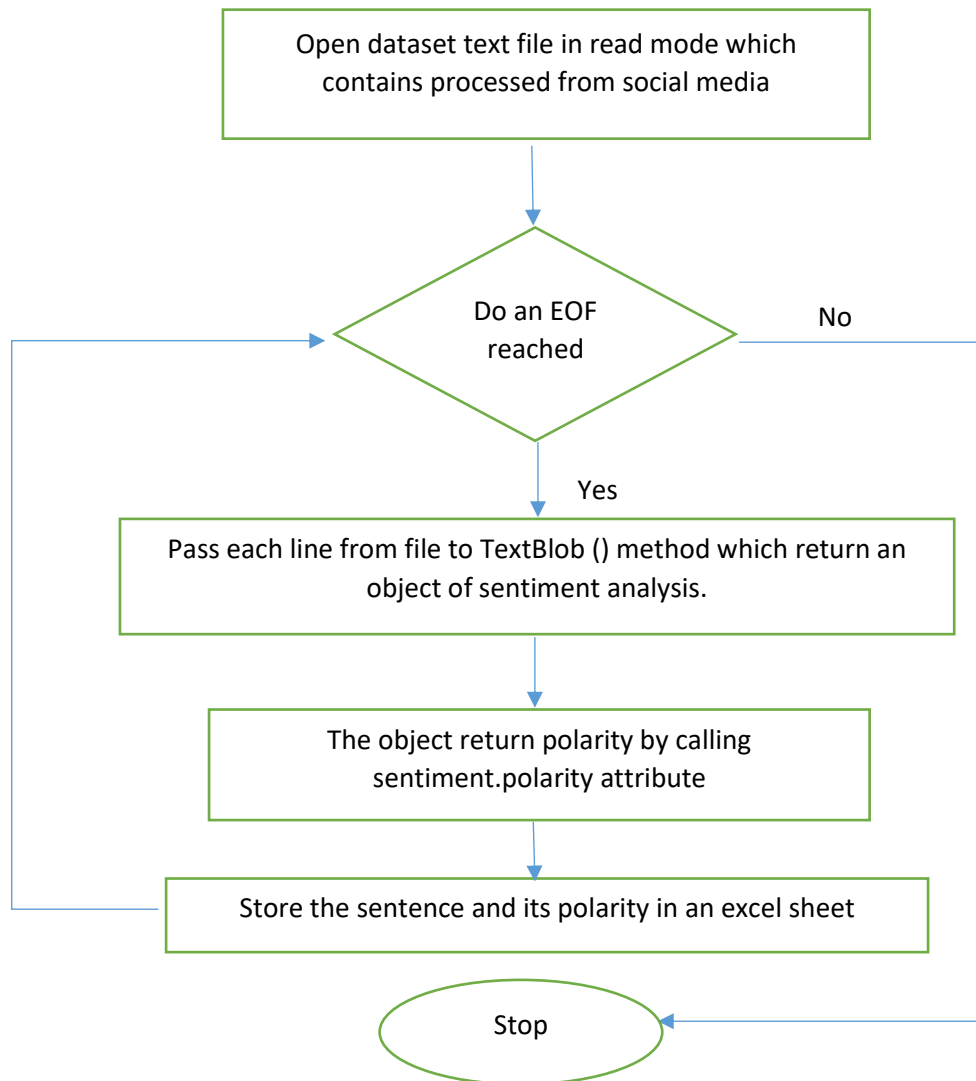
- a) tweepy- tweepy is an open-source Python package that gives you a very convenient way to access the Twitter API with Python
- b) time - provides many ways of representing time in code, such as objects, numbers, and strings

Flowchart



2 b) Data extraction from Google maps:

- Open concerned park's Google Map page
- collected manually from reviews with in required time period
- store in excel sheet for further processing



3. Combine all reviews of twitter and google maps of a park year wise and store as year wise opinions. The total is now categorised into four years i.e., 2019, 2020, 2021 and 2022

4. Find polarity of each sentence from year wise dataset.

The following libraries are used in finding the polarity

a) TextBlob - it is a python library for Natural Language Processing (NLP). It provides a simple API for diving into common natural language processing (NLP) tasks such as part-of-speech tagging, noun phrase extraction, sentiment analysis, classification, translation, and more.

Text Blob returns polarity and subjectivity of a sentence. Polarity lies between [-1,1], -1 defines a negative sentiment and 1 defines a positive sentiment.

b) csv – it helps in various operations on csv files

It is implemented in the following scenario

5. Categorize sentences according to polarity:

Three types of polarities are considering in differentiating the sentences i.e., positive, neutral and negative.

- The excel sheet generated from above flow chart procedure is sorted according polarity
- Store each type of polarity sentences in separate file for find parts of speech.
- This process is applied for all four years processed dataset.

6. Find part of speech for every word in all sentences category wise.

Consider each sentence of a category (positive, neutral or negative) and find parts of speech for every word ‘

Following library packages are used in implementing

a) nltk - NLTK is a library which gives an easy access to a lot of lexical resources and allows users to work with categorization, classification. It is one of the most used libraries for natural language processing and computational linguistics.

b) TreebankWordtokenizer: The Treebank tokenizer uses regular expressions to tokenize text

This tokenizer performs the following steps:

- split standard contractions, e.g. don't -> do n't and they'll -> they 'll
- treat most punctuation characters as separate tokens
- split off commas and single quotes, when followed by whitespace
- separate periods that appear at the end of line

The library provides predated tags for every parts of speech like

VERB — verbs (all tenses and modes)

NOUN — nouns (common and proper)

PRON — pronouns

ADJ — adjectives

ADV — adverbs

ADP — adpositions (prepositions and postpositions)

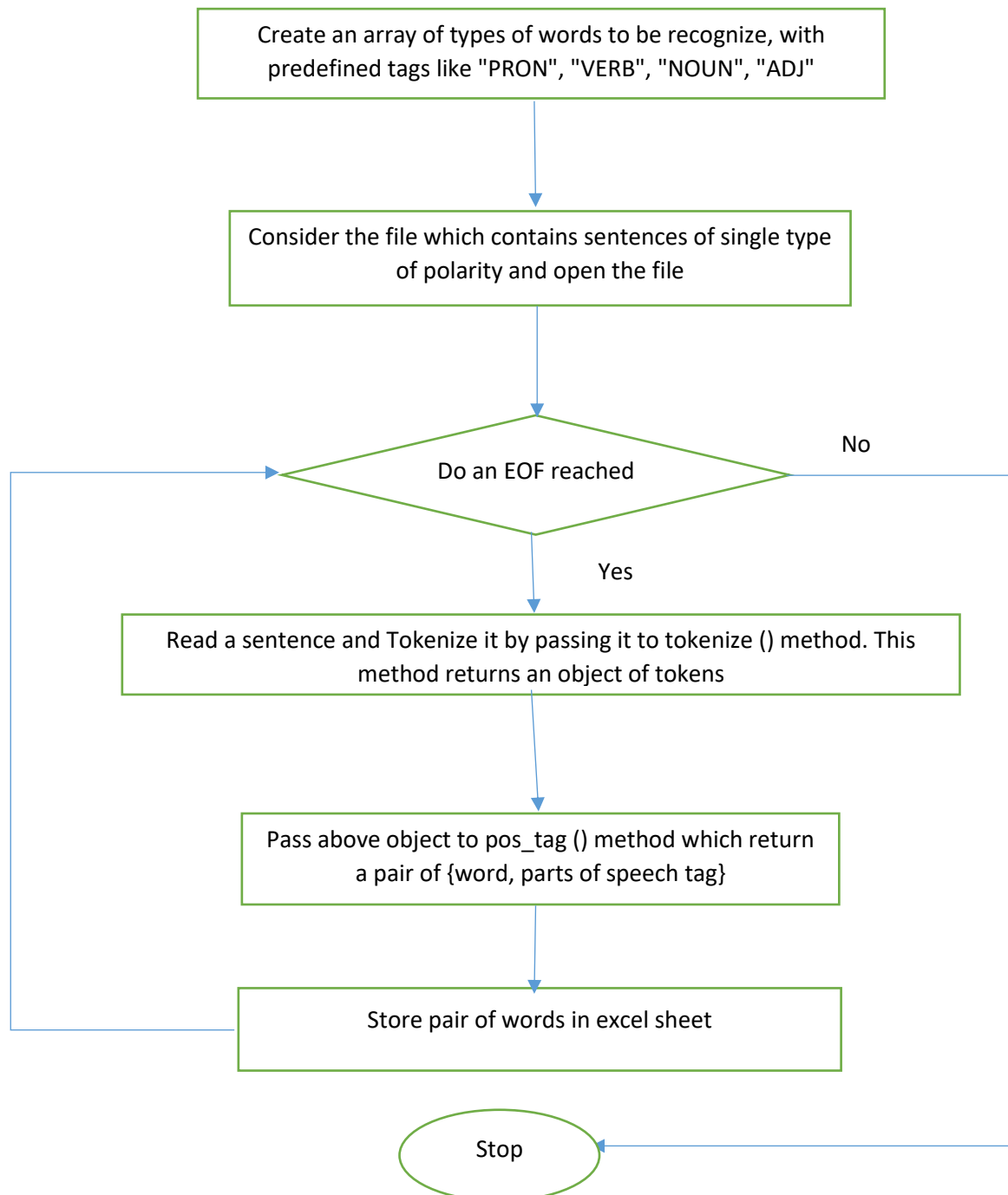
CONJ — conjunctions

DET — determiners

NUM — cardinal numbers

PRT — particles or other function words

The procedure to implement parts of speech on above processed sentences is as follows.



Now the program can generate parts of speech for every word of all sentences in each category of polarity.

7. Collect all verbs and nouns from positive polarity sentences and find word frequencies and word clouds year wise.

The project uses

Steps to implement word clouds and word frequencies

1. Install MAXQDA Analytical Pro latest licensed version
2. Go to the **Import** tab and click on the appropriate icon for your data type, e.g. **Transcripts** to import interview transcripts and **Texts, PDFs, Tables** to import Word files, PDF documents or Excel files.
3. A dialog window will appear, in which you can select the files to be imported. Press **Ctrl** (Windows) or **⌘** key (macOS) to select multiple files at the same time.

Generating a word cloud

Word clouds can be generated for individual documents, document groups, document sets, or for all the documents in a project.

Right-click on the level in the "Document Browser" for which you want to generate the word cloud and select the entry **Word Cloud**.

Alternatively, you can also open the Word Cloud function via the icon with the same name in the **Visual Tools** tab.

A dialog box will then open in which you can select the relevant documents. So that your word cloud does not consist mainly of insignificant words or words bearing very little meaning, you can create a stop list of these words.

The easiest way to do this is to open the **Word frequencies** function in the **Start** tab. The table that appears will list all the words in the text in their order of frequency.

Double-clicking on the green icon in front of a specific word will put it in the stop list, which means it won't be considered for word clouds.

You can also select several words in the list (using Ctrl + or cmd + the left mouse button or Shift + the left mouse button for an area in the list) and then right-click and select **Add to Stop List** to transfer them to the stop list. An icon with the same name is also available in the toolbar.

When you click the **Refresh and apply stop list**, the word cloud is regenerated excluding words from the stop list.

The option **Edit stop list** displays the complete stop list used for word clouds in the current project. You can remove words from this list or add new words..

Working with the word frequency list Via **Start > Word frequencies** you can open a list of all the words contained in the analyzed texts. This list shows, amongst other things, how often each individual word occurs. It is also possible to combine several words to one. To do this, click and drag one or more words onto another word with your mouse

8. Create project website

This website designed with ArcGIS Experience builder licensed software. It takes care of background work of website code building and hosting. We need to concentrate on the way of presentation instead writing programming code. It is a tool which helps to drag and drop the required presentation formats of the data. It takes care of database for the website by uploading the data into ArcGIS from various formats like excel. It is design in such a way that the database, location map of the data and visual representation are interconnected.

Project duration in months : 12

Research Outcomes

- Dashboard on following results
- Distribution of smell categories in urban cities
 - Pairwise correlations between presence of smell categories at street level
 - Heatmaps of Smell categories that are not dominant but are localized in specific areas

Any other details in support of the proposal

Budget

Consumables	Contingency	Travel(field Work)	Others	Total
150000	100000	150000	100000	500000

Grant Application Movement

Name	Designation	College	EMP ID	Date/Time	Action
Dr. Rama Sastry Vedala	Director	HYD - GST - DIRECTORS OFFICE	600723	Aug. 22, 2024, 10:42 a.m.	Recommended
Dr. Mahaboob Basha Shaik	Professor	HYD - GST - CSE	600726	July 31, 2024, 4:19 p.m.	Recommended
Rakesh Kumar Donthi	Assistant Professo r	HYD - GST - CSE	600818	June 8, 2024, 1:35 p.m.	Application Submitted

Comments

Dr. Rama Sastry Vedala|Director|HYD|GST|DIRECTORS OFFICE An amount of Rs. 3.5 Lakhs is recommended, as the work is totally computer based work with some data collection, for which Faculty already made provision under Travel. Contingencies and others have been merged while providing grant.
Aug. 22, 2024, 10:42 a.m. Status :Recommended

Dr. Mahaboob Basha Shaik|Professor|HYD|GST|CSE Appears to be good idea
July 31, 2024, 4:19 p.m. Status :Recommended

Rakesh Kumar Donthi|Assistant Professor|HYD|GST|CSE
June 8, 2024, 1:35 p.m. Status :Application Submitted