

**GOVERNMENT OF INDIA**

**INDIAN INSTITUTE OF REMOTE SENSING**

**ISRO**

Project Report

On

**Landslide Mapping using Geospatial Workflow Technologies**

Submitted by:

|  |  |  |
| --- | --- | --- |
| **Group Members** | | |
| **Name** | **Course** | **Department** |
| Lt Col Ishan Sharma | M.Tech | GID |
| Maj Suraj Choubey | GID |
| Mr Ishan Goyal | GSD |
| Miss Nikita Supekar | NHDRM |

**Git Hub Link**: https://github.com/Nikita-Supekar/landslide\_mod3.2a.git

**TABLE OF CONTENTS**

1. Introduction………………………………………………………………………………...1
2. Objectives…………………………………………………………………………………..1
3. Methodology……………………………………………………………………………..2-8

3.1 Data source……………………………………………………………………………..2

3.2 Business Process Model and Notation Diagram……………………………………….2

3.3 Data extraction using Dashboard…………………………………………………….3-6

3.4 Data extraction using social media…………………………………………………..6-8

3.5 Data Visualisation……………...……………………………………………………....8

1. Limitation………………………………………………………………………………..8-9
2. Future Scope…………………………………………………………………………….…9
3. Conclusion…………………………………………………………………………………9

**Landslide Mapping using Geospatial Workflow Technologies**

1. **Introduction**.

The existing Uttarakhand disaster management portal is an online platform maintained by the Uttarakhand State Govt. It acts as a comprehensive repository which stores all information related to various natural disasters that occur within the state. The aim behind the formulation of the existing website was to enhance coordination among various agencies, departments and stakeholders which are involved in disaster management by providing adequate information to all concerned from one central platform.

The state of Uttarakhand is located in the lap of Himalayas and is prone to various natural hazards such as Earthquake, Landslide, Floods, Forest Fires, etc. The existing portal is a vital digital platform which provides crucial information over a period. However, it has been observed that the existing portal displays only statistical data and lacks substantive photographic evidence. The photographic description helps in easy assimilation and better understanding about the severity and collateral damage caused by the natural calamities.

The aim of using Geospatial Workflow Technology is to extract photographs of landslide/disaster events alongwith the Geo-tagged location details. The details will be stored in the database and can be extracted as and when required. The availability of statistical data along with photographs will enhance interpretation of the severity and consequence caused by natural calamities.

The photographs of disaster events will be collected through crowdsourcing by two methods. One method is by using a dashboard wherein anyone can upload the photographs alongwith Geo-tagged location details. The details will be further stored in the database and can be extracted as and when required. In the second method, social media handles/ web portals including Facebook/ Instagram will be used to extract photographs alongwith location details. The details will be again stored in the same database and can be extracted as and when required. The data from the database can be visualized using various applications like ArcGIS/ QGIS.

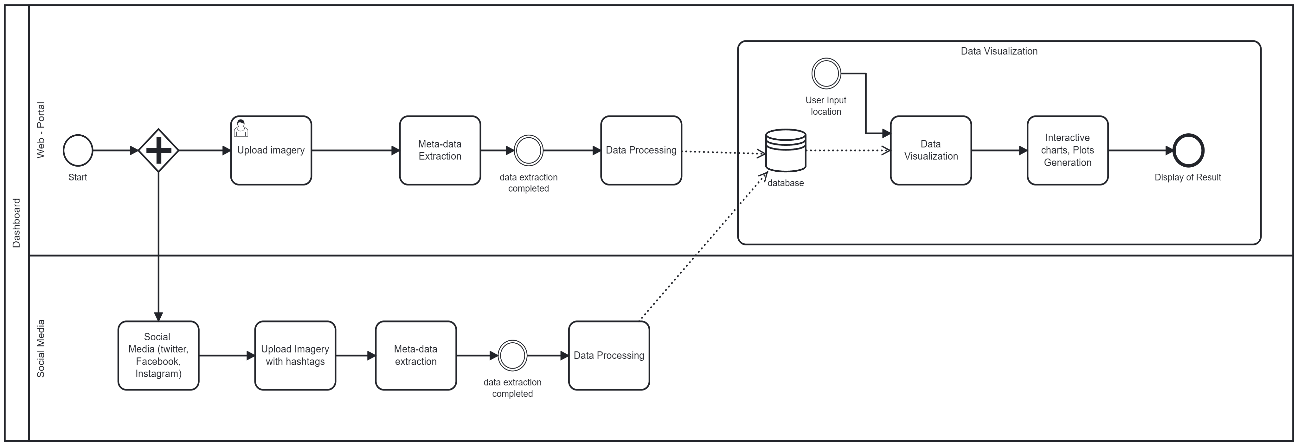
2. **Objectives**.

1. To create a dashboard wherein a user can post real-time photographs related to landslides and other natural disasters and the same can be stored in the database for future reference.
2. To augment the created database by extraction of data from Social Media Platforms (Facebook/ Instagram) through crowdsourcing and adding up to the existing database.
3. To visualise photographs which are stored in database based on the user query for a particular Geographic Location.

3. **Methodology**.

* 1. **Data Sources**. The data will be extracted using two methods as explained below: -

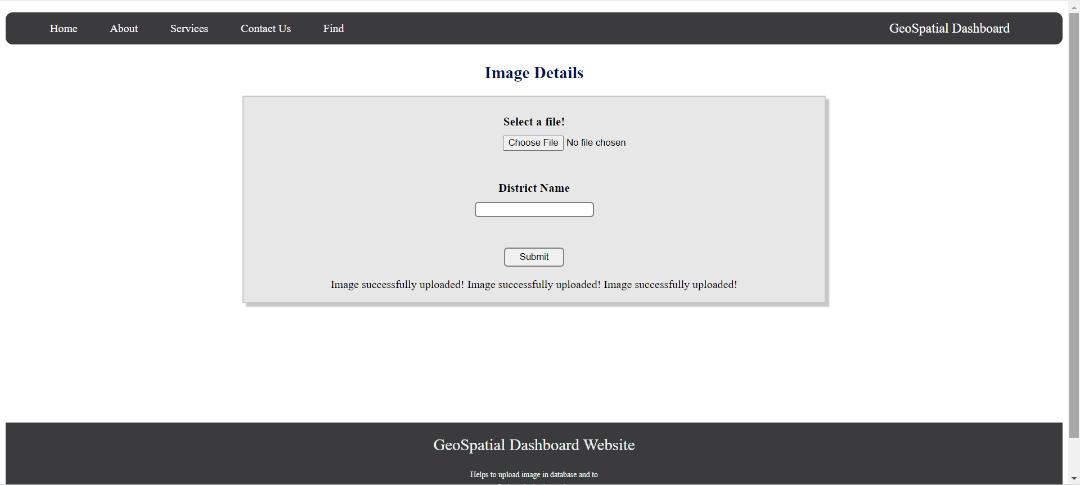
1. User Input Data. A portal will be created on which users can post original photos which are Geo-tagged. The metadata can be further extracted to create Geospatial Database with exact Geo-location and photographic evidence for future reference.
2. Data from Web Portals/ Social Media Handles. In the second method, the dataset to be obtained from Crowd Sourcing (Social Media Handles/ Web Portals). The post containing keywords like Hashtags (For ex: #Chamoli landslides, **#**landslides Dehradun etc.) which have been posted on Social Media Handles (Facebook or Twitter) can be identified and then required data can be extracted to compile the metadata from the photos which will augment the existing database.
   1. **Business Process Model and Notation Diagram (BPMN Diagram)**. The BPMN diagram is used for modelling business processes using graphical elements. It is generally used to represent the flow of activities, events, decisions and gateways within a business process. The BPMN diagram formulated for this project is as given below alongwith explanation.



*Fig 1: Methodology*

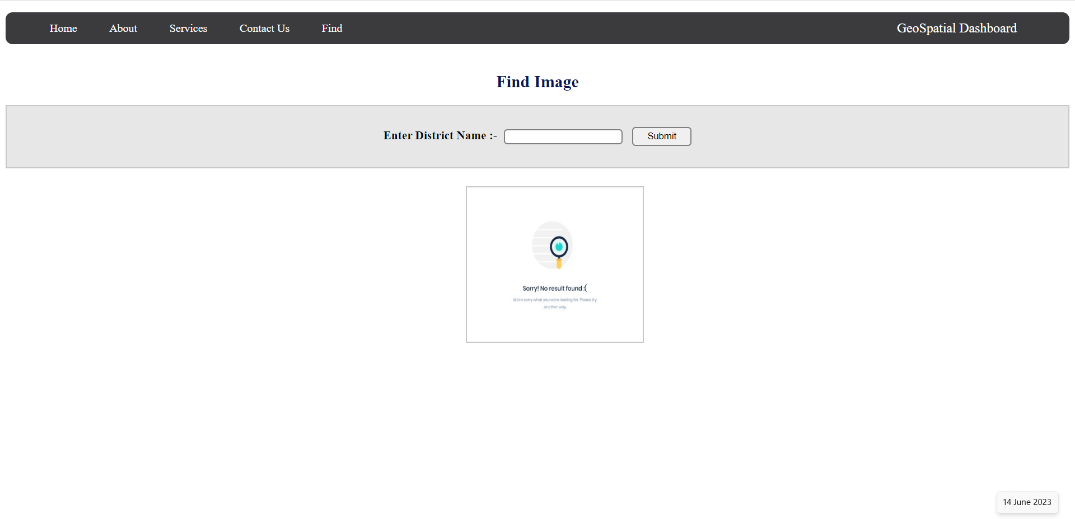
The above BPMN (*Fig 1*) consist of two processes for data extraction. The first process represents a dashboard which will be utilized for two purposes i.e. data inputting and data Visualisation. The data will be extracted using a dashboard wherein anyone can upload the photographs alongwith Geo-tagged location details. The details will be further stored in the database and can be extracted as and when required. In the second process, social media handles/ web portals including Facebook/ Instagram will be used to extract photographs alongwith location details. The details will be again stored in the same database and can be extracted as and when required. Both the processes of data extraction has been explained in proceeding paragraphs.

* 1. First Process (Data extraction using Dashboard). In this process, the end users on the dashboard will upload Disaster data (photographs). Once the photographs, which are Geo-tagged, are uploaded on the portal, the geographic information will be extracted from it and saved in the database alongwith photographs. The screenshot of the dashboard created alongwith brief explanation are given below.



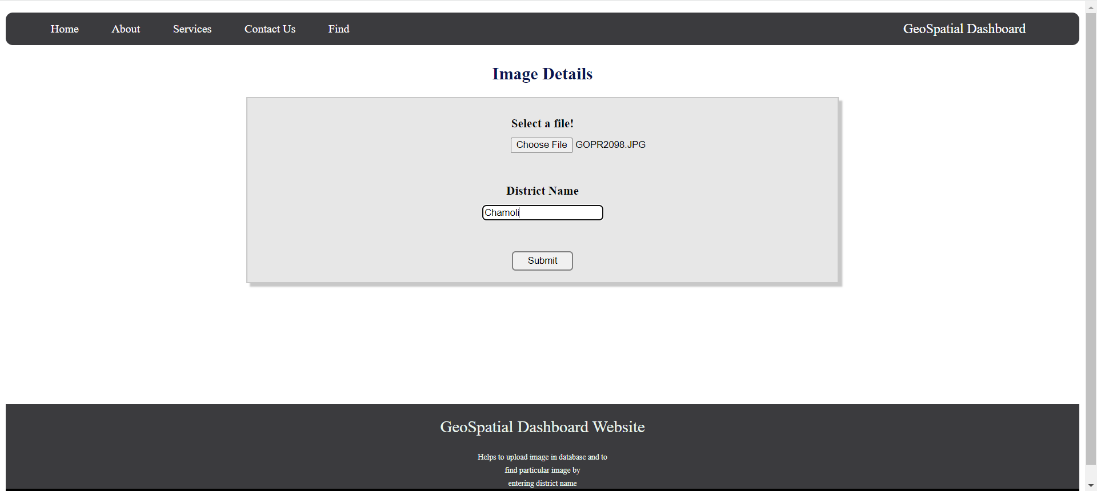
*Fig 2: Geospatial Dashboard*

1. Figure 2 represents Home page of the dashboard created with two important functionalities. The home tab present in the tool box will be used to upload the images. While uploading the image on the dashboard user can give a specific keyword that is the name of the district to which the photograph belongs. The find tab will be used to visualise photographs which will be accessed through keywords (district name) which shall be same as entered while uploading the photographs.



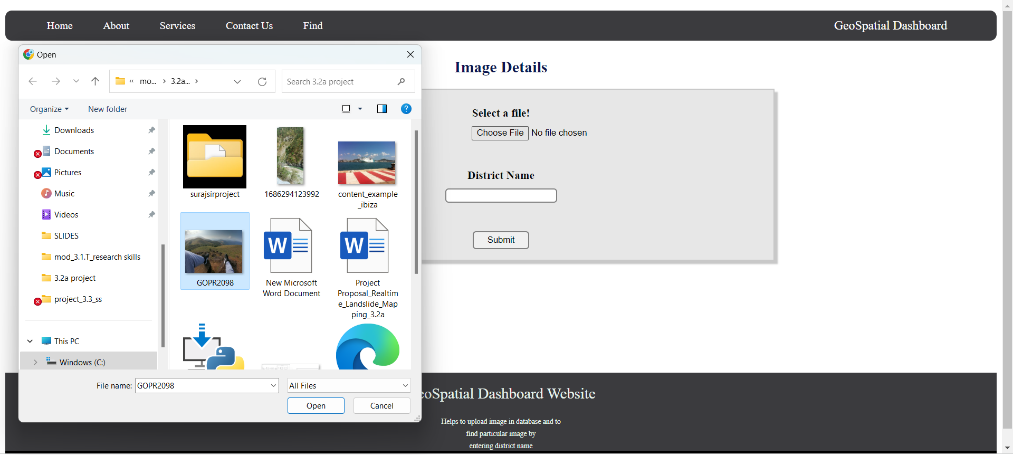
*Fig 3: Data Visualisation*

1. Figure 3 represents visualisation section of the dashboard. In the find tab, present in the tool box, a user can generate a query by providing the name of the specific district. As mentioned above, the district name will act as identifier and will display all the images stored in database related to that particular district.



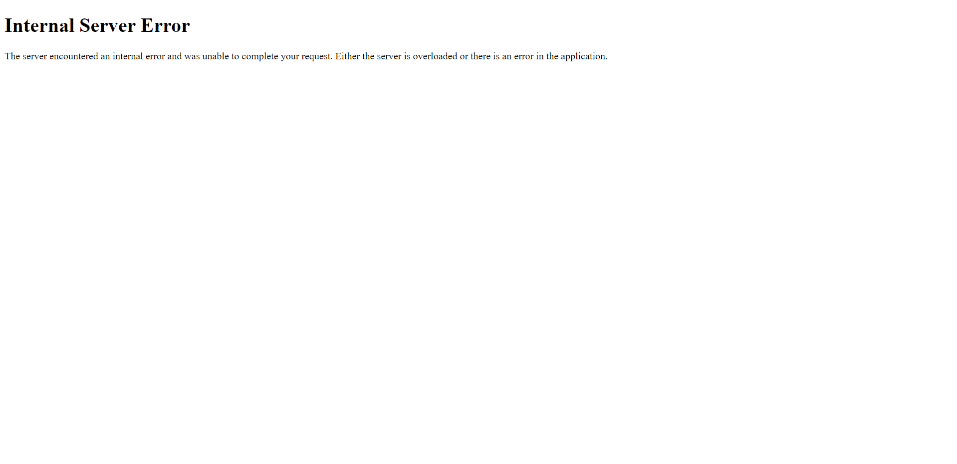
*Fig 4: Data Inputting*

1. Figure 4 represents the steps to upload an image on the dashboard. As created dashboard is a prototype, only one image can be uploaded at one time. In future, this functionality can be improved by having an additional feature of uploading multiple photographs simultaneously at a particular instant.



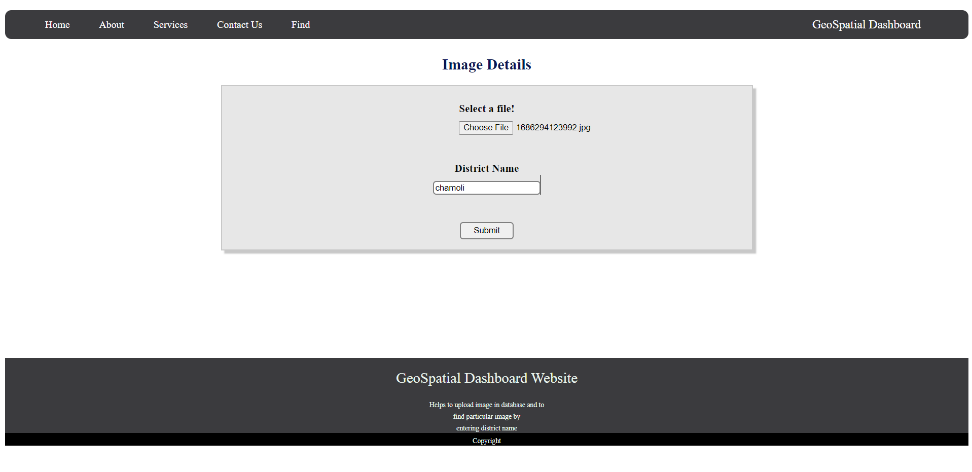
*Fig 5: Image Data Inputting*

1. Figure 5 is used to represent an important step which every end user should ensure before uploading the photograph. All photographs being uploaded must contain geographic location in its metadata and the end user must provide correct keywords (District name) while uploading the photos.
2. Figure 6 represents the error message displayed to the end user in case of uploading an image without geographic location attached with its metadata. This is one of the key functionalities of the dashboard where it is mandatory for an end user to upload only Geotagged photographs with correct district name.



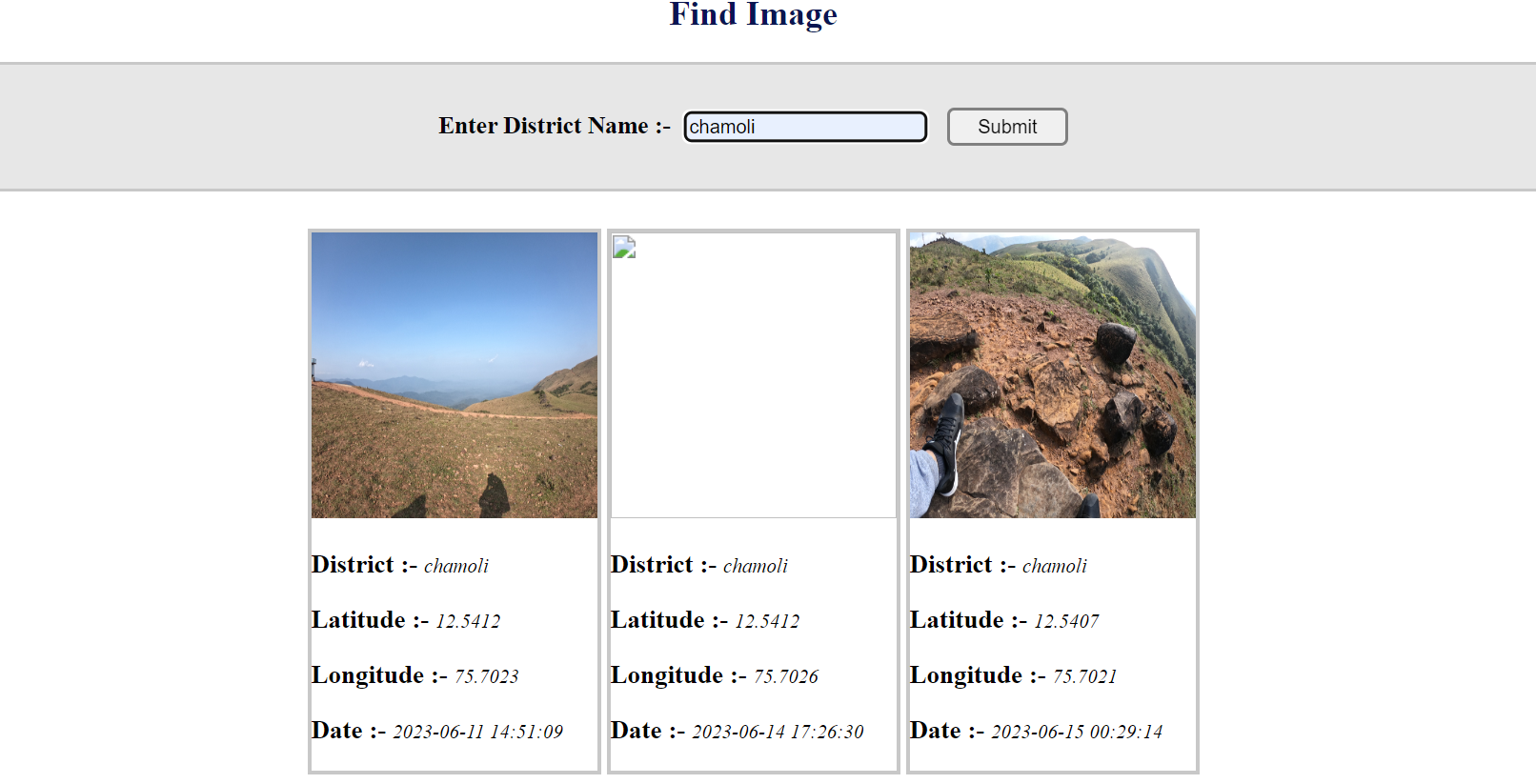
*Fig 6: Invalid Image Uploading Error*

1. Figure 7 represents the correct method of uploading an image. The image being uploaded contains Geo-tag and a correct keyword is also provided by the end user. In this case image will be successfully uploaded and saved in the database.



*Fig 7: Image Metadata Inputtting*

1. Figure 8 represents the visualization section of the dashboard. When an end user creates a query to visualise a photograph of a particular district of Uttarakhand, the backend part of the dashboard retrieves the information from the database and display all the photographs of landslides stored in the database. The visualisation can contain one or more than one photographs depending on the number of photographs stored in database. In case, the generated query has no match, then a message will be displayed stating no image found.



*Fig 8: Image Visualisation*

1. Figure 9 represents the database schema in which the extracted information is stored. For this project, we have used PHP My Admin database where in, a table is created with name Geospatial. The created table being used in this project contains five columns as given below: -

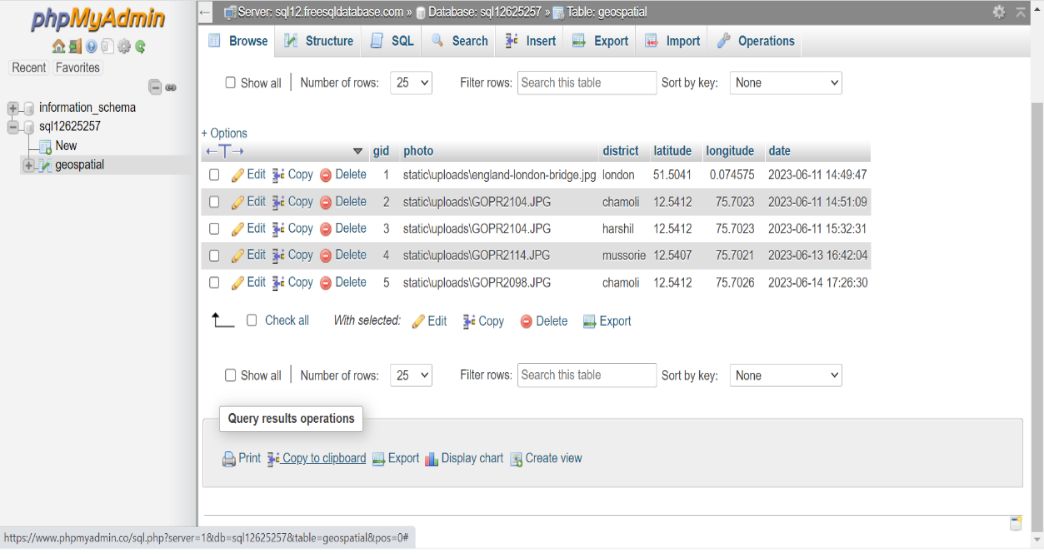
1. GID. It represents a unique ID allotted to each photograph being stored in database.

(ii) District. It acts as a primary key to correctly identify the queried data by user.

(iii) Latitude. This information contains Geographical location only.

(iv) Longitude. This information contains Geographical location only

(v) Date/ Time stamp. An additional information that is the date/ time stamp can also be stored in database.



*Fig 9: Database*

* 1. Second Process (Data extraction using Social Media Handles/ Web Portals: In this phase of data extraction our focus is to extract data from social media platforms such as Facebook, Instagram, and Twitter. These platforms serve as valuable sources of information containing photographic evidence. Whenever a photograph is captured by any device, a location tag is automatically gets attached to its metadata which remain integrated throughout its journey on any social media platform. Our primary objective is to extract these GPS (Global Positioning System) locations from these photographs which have been posted by social media users.

To achieve this, Python (Versatile programming language) is used which is known for its extensive libraries and robust capabilities in handling data. By using Python's set of tools, we will access and extract the desired GPS location data, enabling us to achieve valuable insights from the geospatial context of the images and augmenting them into the database.

The procedure and steps involved have been represented in the UML Activity diagram as given below.

A picture containing text, screenshot, diagram, font

Description automatically generated

*Fig 10: Activity Diagram*

The activity diagram describes a systematic process to extract geographic location from the posted photographs on any social media platform by using python based programming codes. These codes contain two distinct filtering steps to get relevant results and its stepwise description is given below:

Step 1: To select a photograph that belongs to selected location. Python filters hashtags with desired "keywords containing location". This step aims to identify hashtags specifically associated with a particular location. If a hashtag does not contain the required information, the loop will iteratively repeat until a suitable hashtag is found. For our project purpose the keyword considered as “Uttarakhand”. Once the code successfully filters hashtags containing Uttarakhand, it proceeds to the next stage.

Step 2: To identify predefined keywords related to landslide. This stage focuses on identification of hashtags containing information other than location related to landslides and similar terms. By using this filter, the code extracts all relevant hashtags related to landslide.

Step 3: Database creation and storage. After identification of a hashtag that satisfies both filtering criteria, the code will download the associated photograph in each loop. After identification of a hashtag that satisfies both filtering criteria, the code will download the associated photograph in each loop and simultaneously extract the Geolocation from it and stores in the same database with a unique ID.

Python code explanation: The process journey starts with the importation of the image, followed by checking that the photograph contains GPS values or not. If GPS information is found, the process advances towards the extraction of the GPS metadata. However, if GPS data is absent, a "VALUE ERROR" is returned, and the loop starts for another image.

To simplify extracting process, our code depends on the presence of two dictionaries. The first dictionary plays an important role in mapping the EXIF (Exchangeable Image File Format) tags. This mapping makes sure that specific tags within the image's metadata can be accurately identified and accessed when needed. The second dictionary, is to handle GPS tags. By the utilization of these dictionaries, our code navigates through the structure of the image.

Once GPS information is detected, the code proceeds to extract the subtags that consist of GPS-related data holding valuable information such as latitude, longitude, altitude, and timestamp. Subsequently, added in another dictionary, specifically designed to store and analyze this information for further processing. Now the geographic coordinates captured in the degree, minute, and seconds (DMS) format are transformed into a decimal format. This conversion not only standardize the geographic coordinates but incorporates a consideration for the reference direction (N, S, E, W). Thus the code generates comprehensive and accurate geographic coordinates.

After the transformations, the improved GPS information is sent to the database. Each set of coordinates is given a special ID and stored in the database. This valuable data is now ready to be used for various purposes and can analyzed in many different ways.

3.5 **Data Visualisation**: Once the data is stored in the database, users can access it through the dashboard interface. The dashboard facilitates the users to provide input that is their desired location as a parameter. By simply providing the relevant location information, users are able to visualize photographs and collection of photographs depicting landslides of selected region. These photographs not only visualise on the dashboard but also act as a repository of information obtained from crowdsourcing thus it allows the end user to gain valuable knowledge and a comprehensive understanding of the severity and after effects caused by natural disaster in terms of photographic evidence.

4. **Limitations**.

1. The authenticity of the data obtained from crowdsourcing is based on the procedures followed for data acquisition by the public, hence an ambiguity arises as the data is acquired from multiple public sources which require validation in the form of Ground Truth.
2. The acquired data is heterogeneous in nature (collected and uploaded in different formats) and thus different procedures are required for abstraction of data to aid interoperability.
3. The present project is a prototype with limited functionalities in specifically in terms of database requirements. For this project we are saving the images on the local host while the extracted information is stored in database and accessed through SQL query. The concept can be extended by creating a dedicated database where we can store voluminous data with ease of multiple access.

5. **Future Scope**.

1. The current operations are limited to acquisition of crowdsourced data, which is geotagged imagery from social media platforms. However, this limits the data acquisition procedure due to filtering of all such post, which lack image but consist of other crucial spatial information. Hence, to aid semantic enrichment of geospatial information artificial intelligence coupled with the operational algorithm could help in better understanding and visualization of disasters.
2. The present project is targeted for the state of Uttarakhand which can be further extended for the entire nation. This will facilitate different agencies to infer useful and crucial information by creating a repository of photographs all at one place.
3. **Conclusion**.

Landslide mapping using geospatial workflow technologies is an attempt to address the gaps in available information systems where a mere change in the acquisition of geospatial data in the form of geotagged imagery from social media platforms as well as the user not only helps in identifying the highly susceptible locations of the landslide but also, in the creation of a database which further can aid the concerned authorities in damage assessment and resource allocation thus, supporting Disaster Risk Reduction (DRR).