**5.0Research design and methodology***(Provide clear descriptions of software development model adopted, system design and modules and methodology)*

**5.1 Objective 1: To develop a web-based application for the automatic download of Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images of Mwea-west from their respective hubs/ portals**

A web-based application employing an integration of user-friendly client GUI and server applications will be designed, developed and implemented. The general architecture will consist of three (3) main modules that work seamlessly to address near-real time functionalities: -

* **Frontend:** This will include the client user interface that will enable the end consumer(Farmer, technical, political, economic or any other interested parties) to input/query their service of choice via HTTPs get requests.This will run on the browser and naative Mobile application to enable users to intercat with the web app and or portal by cinsumonig functionalities like read and write as they wish. For example.
  + Which areas best support which crop type for maximum yield?
  + Which areas are currenlty covered by which crop (Classification logic)?
  + What is the general crop health at agiven phenological stage of growth?-Crop yield forecast.
  + How has environmental factors e.g rain been feeding the small scale crops in the recent past?
  + Which region(s) were highly affected by the previous crop pest and diseases(If any)-Crop health diagnosis
  + Which regions are nutritionally succeptilbe to poor yield?-Predictive insights.
* **Backend :** This is where the web services and APIs used by front-end architecture and mobile app resides. This will enable server-side web application logic to integrate with the client-side.

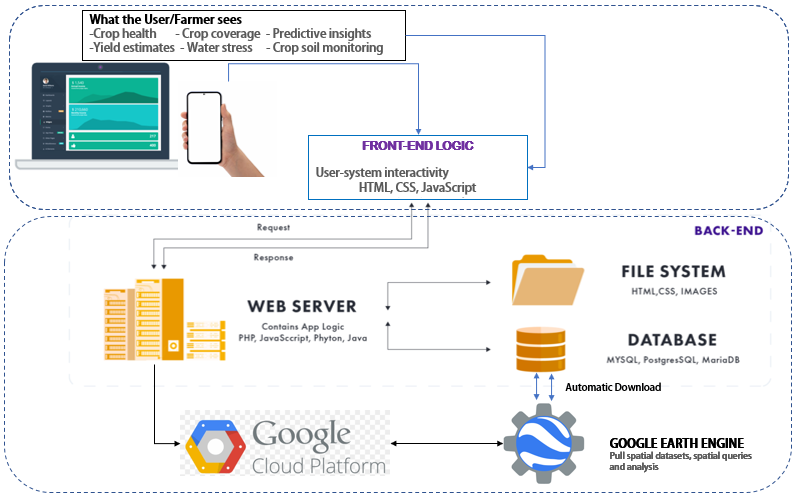
This will be through, the maintenance of the crop mapping central database, making sure that it has high performance and agile responsiveness to requests from the client/user side.

Most of the spatial backend logic will be handled by the Google Earth Engine Cloud platform especially now that it hosts a vast variety of spatial datasets in cloud repository that gets updated on a near daily basis.

* **Full-Time hosting services:** This will make the applications and websites accessible to any user who has access to cloud resources like internet. Will constitute a network of connected virtual and physical cloud servers to host the application and website, ensuring greater flexibility and scalability specially to run the complex machine learning and IA enabled algorithms.

Key Functional features of the proposed hosting services:

* Web and Mobile Applications and solutions are deployed on a cloud network like Google Cloud Platform (GCPs) for easy sink with the GEE backend logic.
* Resources are scaled to user needs as indicated in the front-end logic.
* Querying support using SQL (including PostgreSQL) or NoSQL databases like spatial GeoJSONs.
  + Enables the Small-Scale Crop Mapping Solutions are to run autonomously and controlled using APIs, web portals, and the mobile apps.



*Fig: Web and Mobile based Client-server architecture, to automatically download from GEE query and display results*

**5.2 Objective 2:To implement real-time image processing and archiving for retrieval of the Landsat 8/9, Sentinel-2A/2B, and Sentinel-1A/1B images of Mwea West, as they become available.**

Real-time geoprocessing functionalities is considered key especially when near-time decisions need to be achieved. This will be implemented by running on the server, background image processing algorithms will be running and processing the image data to various levels, as may be requested by the user. The main processing algorithms include:

1. Landsat 8 and Sentinel-2A/2B: At-sensor radiance and TOA reflectance, Surface reflectance, Compositing, Cloud Masking and Mosaicking.
2. Sentinel 1A/1B: Pixel-wise derivation of the backscatter coefficient through application of orbit file, GRD border noise removal, thermal noise removal, radiometric calibration, and terrain correction (orthorectification).
3. The preprocessed image will then be fed into Machine Learning/ AI algorithms like Neural networks, Support Vector Machine Classifiers, Decision trees to perform the AI/ML logic.

5.3 **Objective 2: To develop and implement AI/ ML image classification algorithms for mapping of land cover and final estimation and retrieval of crop area, crop status indices and crop yield.**

During the client data access request process, users will be required to stipulate which kind of data output they wish to view or download. This research will implement automatic AI/ML-driven classification algorithms as implied in 5.2(c) above to generate:

1. Land cover classification (general), depicting all classes as outlined by the UN Land Cover Classification System’s initial ‘Dichotomous Phase’, in which eight major land cover types are defined:

(1) Cultivated and Managed Terrestrial Areas,

(2) Natural and Semi-Natural Terrestrial Vegetation,

(3) Cultivated Aquatic or Regularly Flooded Areas,

(4) Natural and Semi-Natural Aquatic or Regularly Flooded Vegetation,

(5) Areas associated with artificial bare lands

(6) Artificial Waterbodies, 

(7) Snow and Ice, and

(8) Natural Waterbodies, Snow and Ice (Di Gregorio, 2005).

1. Cropland area including Cultivated and Managed Terrestrial Areas and Cultivated Aquatic or Regularly Flooded Area

The classification of the satellite images will require ground reference data collected at

1. Crop-related indices including Normalized Difference Vegetation Index (NDVI), Normalized Difference Red Edge (NDRE) index, Modified Soil Adjusted Vegetation

Index (MSAVI), Normalized Difference Moisture Index (NDMI) and Red edge

Chlorophyll (ReCl) index.

1. Field-scale time series analytics derived from the indices above.

The classification of the satellite images will require ground reference data collected at regular intervals over the course of the research project. JKUAT possesses two state of the art drones (multi-spectral and LiDAR) and RTK-GNSS survey equipment which will be used for the ground truth data acquisition exercises.

