# Optimizing Smart Library Spaces: Integrating PIR Sensors, Credit-based Booking Systems, and Advanced Algorithms for Efficient Resource Management and Space Allocation

Mohamud Rashid Mohamud
Huawei Technologies Co. (Kenya) LTD
P.O Box 66430-00800 Nairobi, Kenya

H. Chege Nganga
Jomo Kenyatta University of Agriculture and Technology
P.O Box 62000 – 00200 Nairobi Kenya

Stephen Kiambi
Jomo Kenyatta University of Agriculture and Technology
P.O Box 62000 – 00200 Nairobi Kenya

## ABSTRACT
A smart library is designed to bring digital intelligence to objects and spaces, enabling real-time data-informed decisions. This aspect is enabled by the use of sensors that are connected to the internet, all collecting and sharing data and, in this case, data about traffic information in the library, enabling users to book seats, pinpoint locations of vacancy, and intentionally manage energy (power) and resources (labor). The benefit of monitoring occupancy data is introducing an efficient communication channel, revolutionizing a traditional library into a data source hub. This allows a smooth interaction with the user, monetization of the service, adjustment of spatial design and traffic flow, and movement to greener spaces. This paper provides an approach to the above advantages by incorporating a hardware system involving Passive Infrared (PIR) sensors, a node Micro-controller Unit (MCU), and a web app. The prototype system allows seat booking, calculation of library credits, and library occupancy prediction using the Random Forest model and Random Forest Regressor to optimize space allocation and resource management.

## General Terms
IoT, Machine Learning, Smart Spaces, Artificial Intelligence, Energy Efficiency, Predictive Analytics, Credits, Micro-controllers.

## Keywords
PIR Sensors, NodeMCU, Random Forest, Random Forest Regressor, Occupancy, Model.

## 1. INTRODUCTION
Libraries have always been regarded as sanctuaries of knowledge. The origin of the earliest libraries dates back to 600 BC when there was a need to maintain written records during the earliest civilization times [1]. Various stakeholders have always focused on developing safe, conducive, and comfortable spaces enabling people to learn, work and enjoy their favorite books.

The main challenges currently facing libraries across the globe include inadequate funding, insufficient seating and study areas, and inadequate infrastructure [2]. The current system is also flawed because of the uncertainty in accessing and managing space, staffing, energy, and revenue collection. This unreliability leads to:

a) Overcrowding: The traffic is not spread out over the hours optimally.
b) Time wastage: Making unnecessary trips to find no space and taking a lot of time to settle down while looking for seats.
c) Poor staffing: The number of librarians needed at a particular time to ensure effective management is not data-based.
d) The company cannot recognize new revenue streams created by introducing marketing prospects.
e) Unacceptable loss of energy in lighting and HVAC technologies.

This proposal aims to solve these problems by inferring the library as a chargeable booking service where users and management enjoy the convenience of a digital system with real-time updates. Users can book a seat beforehand and be assured that space will be available when they arrive. Passive Infrared (PIR) sensors will be installed at the library seating areas to detect human presence. The sensors will continuously monitor user presence, and with the help of a microcontroller, the sensor data is sent to a cloud-based database. Further, the proposal seeks to ensure that energy is efficiently utilized by automatically switching off all equipment when not in use.

Additionally, to achieve efficient resource allocation in the future, predictive modeling algorithms such as Random Forest can be used to predict times of the year, month, or day when the library is busiest. The model's output will be predicted occupancy information that will serve as an input for the optimization model. The optimization model will ensure optimal resource allocation, i.e., workforce required and expected energy consumption. There is estimated to be about 15 -30% reduction in energy consumption and about 20% savings on management and labor expenses by aligning staff schedules and resource allocation with the predicted occupancy [3].

## 2. LITERATURE REVIEW
### 2.1 Introduction
Library buildings typically offer peaceful spaces ideal for individual study and areas designed for group work and collaboration. They may provide public facilities to access their electronic resources, such as computers, desktops, and internet access. Libraries can serve as community centers, offering programs and opportunities for people to participate in lifelong learning. To improve service delivery, it is essential to study building occupancy patterns to improve design solutions and better understand space utilization [4]. Technological advancements like the Internet of Things (IoT), Artificial

Intelligence (AI), Big Data, and Cloud Computing can be utilized to maximize the Return on Investment (ROI) of libraries. IoT technologies can offer comprehensive management of library spaces and resources. Additionally, incorporating AI concepts such as machine learning and big data technologies can improve the library experience by deeply analyzing users' patterns and predicting their study patterns, thus providing personalized services to library users and staff [5].

## 2.2 Internet of Things (IoT)

### *2.2.1 Introduction*

The Internet of Things (IoT) refers to a wide network of physical objects, often called "things," embedded with sensors and other technologies, enabling them to interact and share data with other devices and systems via the Internet (see Figure 1). A 'thing' refers to an item or device not usually connected to the Internet, such as wearables, bulbs, watches, washing machines, etc. These connectable devices use sensors to enable remote monitoring, status, manipulation, and evaluation of trends of such devices [6][7][8].

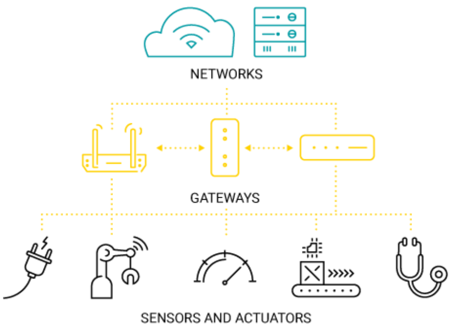


**Figure 1: IoT Ecosystem [10]**

A simple IoT system architecture comprises three layers: Devices, Edge gateway, and the Cloud. Devices include networked things, such as the sensors and actuators in IoT equipment, particularly those that use Bluetooth, Zigbee, Wi-Fi, or proprietary protocols connected to an edge gateway. An edge gateway is a network entry point for devices typically accessing cloud services. The gateways are sensor data aggregation systems that provide functionality, such as pre-processing data, securing connectivity to the Cloud, using systems such as Web Sockets, the event hub, and, in some cases, edge analytics. They also often provide network translation between networks that use different protocols. The endpoint is the cloud application built for IoT, which includes various database systems that store sensor data. The cloud layer in most cloud-based IoT systems facilitates event queuing and messaging systems that handle communication in all layers [9].

The IoT system is implemented to drive efficiency and improve quality of life. Users decide how to utilize the devices, networks, and platforms to achieve these results and realize the full potential of IoT.

### *2.2.2 IoT Based Smart Spaces*

Okoronkwo et al. proposed using Radio Frequency Identification (RFID) and pressure sensors mounted on seats to monitor occupancy. The RFID reader gets details of the seat occupant, which are sent to the web server. Students can remotely log in to their information on the library web application to view vacant seats. In contrast, the Librarian can view occupants' details (on request from the student) [11].

Chang and Kim proposed using thermal, passive infrared, and ultrasonic sensors to monitor seat occupancy. The occupancy data is then relayed to a cloud-based system, whereby users can see library occupancy from a web-based platform [12].

Shokrollahi et al. deeply analyzed the capabilities of PIR sensors when monitoring space occupancy. The sensors can detect even the slightest movements, thus providing valuable insights that can be used for behavioral analysis of people in buildings [13].

The proposals mentioned above have the following limitations that have been addressed in this study:

a) The systems do not provide user time monitoring in the library.
b) Data analytics can improve user experience. The described systems do not provide data insights to users who are unaware of library busy times.
c) It would be ideal if students booked a library seat before their arrival. This would help them save time looking for seats in the library. The discussed systems do not provide booking services
d) Implementing a library credit system would be prudent to ensure that users utilize hours matching the credits they pay for.

## 2.3 ML in Smart Spaces

The three most common machine learning approaches in occupancy prediction are supervised, unsupervised, and reinforcement learning. Supervised learning involves training the model using labeled historical data. In unsupervised learning, the data set is not labeled, and the model is trained to identify patterns, structures, or relationships. With reinforcement learning, an agent learns to make decisions based on its outcome after interacting with the environment. The agent is punished or rewarded based on performance while striving to achieve maximum rewards over time [14].

Wang et al. proposed the use of three models to predict the occupancy of buildings which include; Backpropagation Artificial Neural Networks (ANN), Support Vector Machine (SVM), and k – nearest neighbor algorithm. The study aims to achieve three goals: Firstly, using a fusion framework to develop a data-driven occupancy modeling approach. Secondly, to check which data sets are most preferable for predicting occupancy. Finally, which model (ANN, kNN, and SVM) provides the most accurate prediction [15].

In another publication, Sirmacek and Riveiro used a computer vision-based algorithm together with a feature classification method to predict the occupancy of office spaces. Heat sensors have been used to monitor human presence [16]. Although the computer vision and feature-based algorithms provide 80% and 90% occupancy prediction, the thermal sensors are prone to collecting false positives as they can detect non-human objects, such as humans. Additionally, they do not factor in aspects like energy and resource management prediction.

# 3. METHODOLOGY

## 3.1 Introduction

This section thoroughly elucidates the different components used in the project, from data collection to data storage and, finally, data analysis, visualization, and prediction. It has incorporated both hardware and software components utilized, a detailed description of how the different hardware devices are interfaced with each other, and how those devices communicate with the cloud. Block and circuit diagrams have been used.

## 3.2 Required Hardware, Tools, and Software

a) Google Colab provides an environment to run Python Scripts used for modeling and free computing resources for training models.
b) Fritzing tool: used to develop the network schematics of the hardware, including all the connections.
c) NodeMCU: The node Microcontroller Unit (MCU) is a firmware and development kit used to build IoT applications. It has processing capabilities and can transmit data via its embedded Wi-Fi module. It also obtains data from sensors connected to it.
d) Passive Infrared (PIR) Sensors detect changes in Infrared Radiation (IR) produced by objects, especially animals.
e) PIR sensor data set: Hong et al. provided a room occupancy data set with PIR sensors set in 51 rooms [17].
f) LCD Display: An LCD Display to visualize the occupancy status for users in the library.
g) Channel Relays are used to switch on or off equipment when users are present or absent. This includes lighting and HVAC, thus ensuring efficient energy utilization.

## 3.3 System Architecture

As illustrated in Figure 2, PIR sensors have been mounted on stationary tables. They detect human presence using emitted sound waves with higher frequency than the human ear can perceive. An ESP8266 Wi-Fi module is embedded in the node Microcontroller Unit (nodeMCU). It is used to send the response of the sensors to the cloud periodically, preferably at intervals of five minutes. A 5V single-channel relay switch is also connected to the microcontroller. When a PIR sensor detects no one around there, the node Microcontroller Unit using the relay switch will then switch off all devices associated with that particular PIR sensor. Whenever the PIR sensor detects that a person is seated, the microcontroller signals the relay to switch on the associated device or even lights. An LCD is also used to depict the status of library occupancy in offline mode, which is thus preferable when the network is affected by downtimes.

The cloud platform supports secure sensor data collection, data analysis, and visualization, providing valuable predictive modeling insights. In this case, the PIR sensor data on the occupancy of libraries is sent to a cloud server storage, which is visualized in charts.

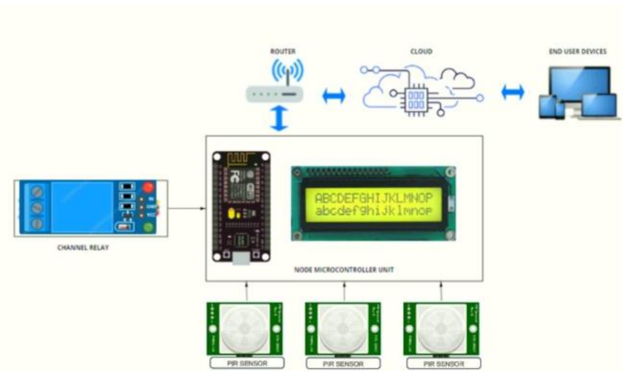


**Figure 2: IoT Ecosystem**

A user books a seat on the library booking platform. For a successful booking, the student must have library credits. Users can also forgo booking, but upon arrival at the library, the librarian will assign them any other available seat that hasn't been booked.

## 3.4 Predictive Modelling

### *3.4.1 Introduction*

Two models (Random Forest Classifier and Random Forest Regressor) were used to predict occupancy and energy usage in kilowatt hours (kWh) over a 12-month period.

The Random Forest algorithm is a supervised learning technique that utilizes multiple decision trees on a given data set to improve its predictive accuracy. One single tree is insufficient to make the final decision; thus, the Random Forest collects the prediction from each tree and predicts the final output based on most prediction votes from the other trees, as shown in Figure 3 below [18].

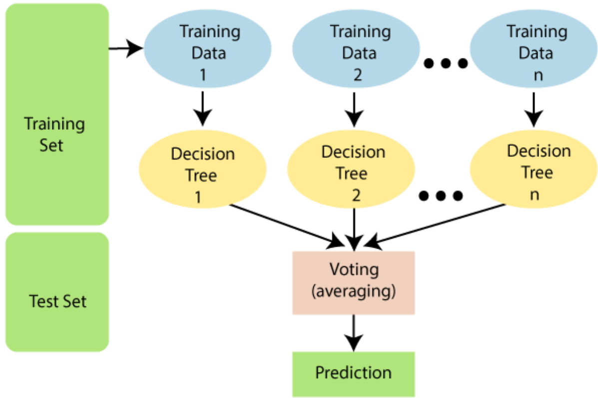


**Figure 3: Random Forest Algorithm [18]**

The Random Forest Classifier model is used to solve classification issues, such as predicting whether the library is occupied. The Random Forest Regressor algorithm, on the other hand, is used to solve regression issues, such as predicting a continuous numerical value like energy consumption.

### *3.4.2 Prediction Methodology*

The following steps were used to predict library occupancy and estimate energy consumption successfully:

1. **Data Collection and Preparation**
   The dataset containing library occupancy and energy usage data was obtained. The dataset included features such as date, time, day of the week, and occupancy status. The data was cleaned to handle missing values, outliers, and inconsistencies. Public holidays, exam periods, and university closures were accounted for, and irrelevant data points were filtered out. Feature engineering was applied to derive additional attributes such as hour, day of the week, and periodic transformations (cosine and sine) of the hour for capturing cyclical patterns.
2. **Feature Engineering**
   The following features were engineered to improve the model's ability to capture trends and patterns:
   a) **Hour of the Day**: Captures variations based on time.
   b) **Day of the Week**: Identifies differences in behavior across weekdays and weekends.

   c) **Weekend Indicator**: Distinguishes weekends from weekdays.
   d) **Periodic Transformations**: Sine and cosine transformations of the hour to account for cyclical patterns in time.

3. **Occupancy Prediction Using Random Forest Classifier**
   A **Random Forest Classifier** was trained to predict occupancy status (0 or 1) based on the engineered features. The steps were:
   a) Splitting the dataset into training and testing sets (80:20 ratio).
   b) Performing hyperparameter tuning using grid search with 5-fold cross-validation to identify the best parameters.
   c) Training the optimized Random Forest Classifier on the training data.
   d) Evaluating the model on the test set using accuracy as the primary metric.
4. **Energy Prediction Using Random Forest Regressor**
   A **Random Forest Regressor** was trained to predict energy consumption (in kilowatt-hours) based on occupancy and other features. The steps were:
   a) Predictive occupancy and other engineered features are used as additional input features.
   b) Splitting the dataset into training and testing sets (80:20 ratio).
   c) Training the Random Forest Regressor on the training data with pre-determined hyperparameters.
   d) Evaluating the regressor using Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE).
5. **Monthly Cumulative Metrics**
   Predictions for both occupancy and energy usage were aggregated every month:
   a) **Cumulative Occupancy**: The rolling sum of predicted and actual occupancy was calculated.
   b) **Monthly Energy Usage**: Total predicted and actual energy usage were computed monthly.
6. **Visualization of Results**
   The results were visualized to compare actual vs. predicted values:
   a) A plot of **Monthly Energy Usage** for actual and predicted values.
   b) A plot of **Monthly Cumulative Occupancy** for actual and predicted values.

# 4. RESULTS DISCUSSION
## 4.1 IoT Prototype
The PIR sensors successfully detected human presence, and the occupancy data was relayed to a cloud-based database. Additionally, the user could see the available and occupied seats in the library. The user could also reserve a seat in the library, and the credits were deducted or added accordingly. Figures 4 and 5 represent the hardware prototype in action, real-time occupancy status, and the credit system. The channel relay system could switch on or off associated equipment based on the occupancy status of the PIR sensors.

**Figure 4: Complete prototype**

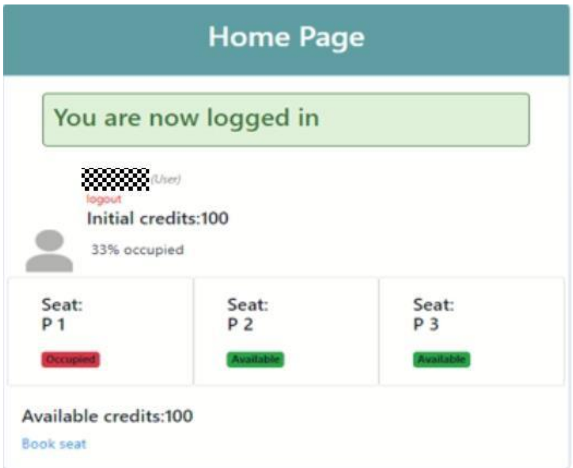


**Figure 5: Occupancy status and available user credits**

## 4.2 Model Result
The study demonstrates a robust methodology for predicting library occupancy and energy usage based on historical data. The occupancy prediction model, developed using a Random Forest Classifier, achieved an accuracy of 87.65% in classifying occupancy states (occupied versus unoccupied), as shown in Figure 6 below. This high level of accuracy highlights the effectiveness of the feature engineering approach, which included temporal features such as hour of the day and day of the week, as well as sinusoidal transformations to capture cyclic variations. The model also accounted for data imbalances using class weighting during hyperparameter tuning, further enhancing prediction reliability.

For energy consumption prediction, as shown in Figure 7, a Random Forest Regressor was employed, resulting in a Mean Absolute Error (MAE) of 0.032 kilowatt hour (kWh) and a Root Mean Square Error (RMSE) of 0.041 kWh on the test dataset. These metrics indicate that the model could closely estimate energy usage based on predicted occupancy and other temporal features. When evaluated at the monthly aggregation level, the expected energy usage values aligned strongly with actual energy consumption trends, affirming the model's ability to capture broader usage patterns.

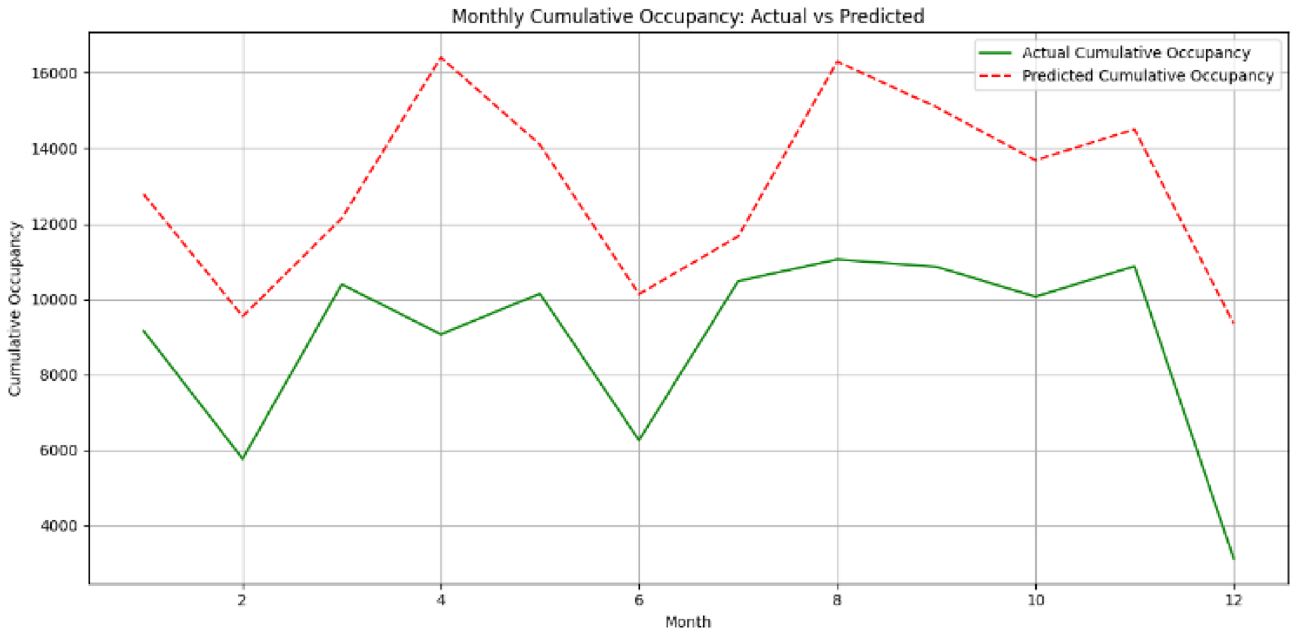


**Figure 6: Actual vs Predicted Occupancy**

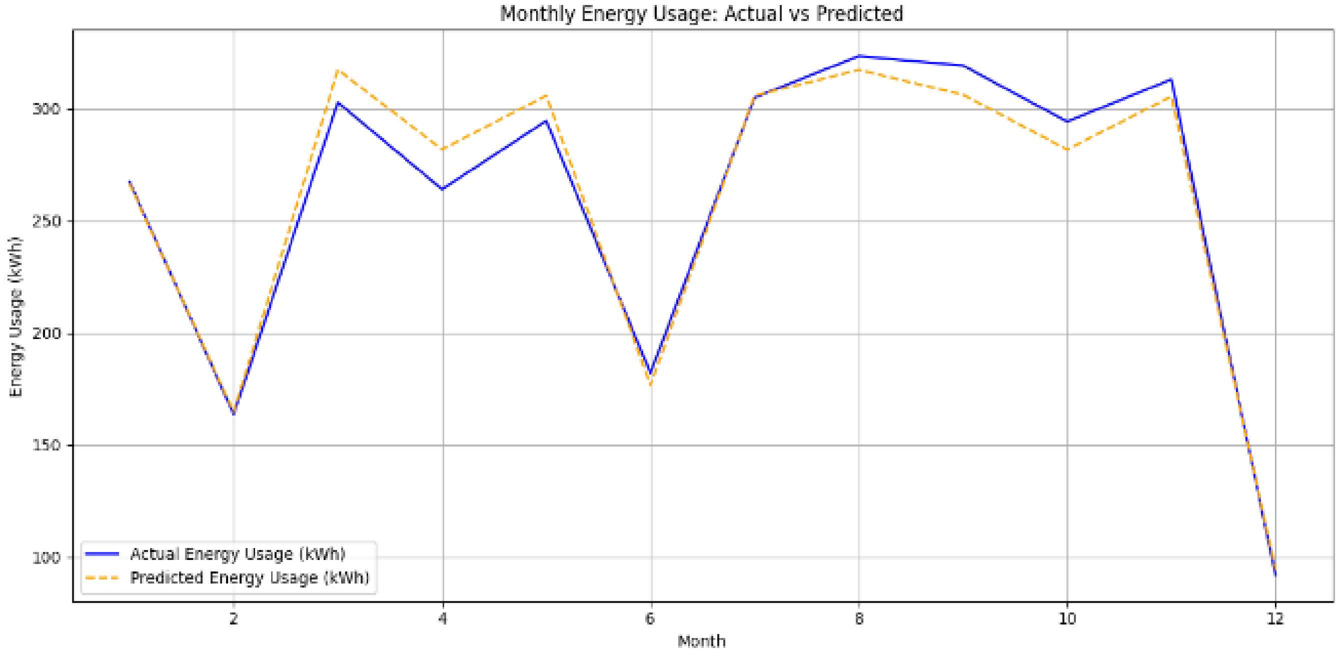


**Figure 7: Actual vs Predicted Occupancy**

## 5. CONCLUSION

This study presented an efficient way to run and access libraries. The design incorporates PIR sensors that provide a current percentage of seats occupied. The website portal presents a simple GUI of this occupancy and booking functionality for the client. It further incorporated predictive modeling of occupancy and energy consumption using a random forest algorithm with an average accuracy of 87.65%. The results underscore the feasibility of applying machine learning techniques for occupancy and energy management in dynamic environments. The integration of predictive analytics can facilitate more informed decision-making, ultimately contributing to enhanced energy efficiency and operational planning. In the future, gradient boosting algorithms can be used to further improve performance. Finally, optimizations can be done on the overall design to learn user preferences and further enhance the user experience in the library.

## 6. ACKNOWLEDGMENTS

Gratitude is extended to Jomo Kenyatta University of Agriculture and Technology (JKUAT) for providing the opportunity to work on this project. Sincere appreciation is also expressed for the invaluable support of Mr. Chege Ngángá and Mr. Stephen Kiambi throughout this research.

## 7. REFERENCES

[1] Adaora Joy Udo-Anyanwu, “CHAPTER ONE ORIGIN OF LIBRARIES,” ResearchGate, pp. 1–22, Nov. 2021, Available:https://www.researchgate.net/publication/355942072_CHAPTER_ONE_ORIGIN_OF_LIBRARIES

[2] A. Ullah, M. Usman, and M. K. Khan, "Challenges in Delivering Modern Library Services in the 21st Century," International Journal of Social Science Exceptional Research, vol. 2, no. 6, pp. 146–151, Jan. 2023, doi: https://doi.org/10.54660/ijsser.2023.2.6.146-151

[3] N. Haidar, Nouredine Tamani, Yacine Ghamri-Doudane, and Alain Bouju, "Occupant Behavior Prediction and Real-Time Correction-based Smart Building Energy Optimization," GLOBECOM 2022 - 2022 IEEE Global Communications Conference, pp. 1–6, Dec. 2020, doi: https://doi.org/10.1109/globecom42002.2020.9348056.

[4] Q. Wang, H. Patel, and L. Shao, "A longitudinal study of the occupancy patterns of a university library building using thermal imaging analysis," Intelligent Buildings International, vol. 15, no. 2, pp. 62–77, Nov. 2022, doi: https://doi.org/10.1080/17508975.2022.2147129.

[5] P. Liu, "Design and Implementation of Library Seating Management System," Journal of Computer and Communications, vol. 12, no. 08, pp. 292–306, Jan. 2024, doi: https://doi.org/10.4236/jcc.2024.128018.

[6] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Ayyash, S. A. Shah, and D. D. S. P. P., "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2347-2376, Fourth Quarter 2015. doi: 10.1109/COMST.2015.2444095.

[7] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions," Future Generation Computer Systems, vol. 29, no. 7, pp. 1645-1660, Sep. 2013. doi: 10.1016/j.future.2013.01.010.

[8] A. Zanella, N. Bui, A. Castellani, L. V. S. Pellegrini, and M. Zorzi, "Internet of Things: A Survey on Technologies, Protocols, and Applications," Computer Networks, vol. 56, no. 15, pp. 2787-2805, Oct. 2014. doi: 10.1016/j.comnet.2014.05.014.

[9] Hassan, Qusay; Khan, Atta; Madani, Sajjad (2018). Internet of Things: Challenges, Advances, and Applications. Boca Raton, Florida: CRC Press. p. 198.

[10] "IoT ecosystem: 4 key elements" [Online]. Available: https://www.avsystem.com/blog/iot-ecosystem/

[11] Okoronkwo et al, "Smart Library Seat, Occupant and Occupancy Information System, using Pressure and RFID Sensors. 10.1109/NEXTCOMP.2019.8883610.

[12] D. Chang and M. Kim, "Library Occupancy Sensor," Jan. 2020.

[13] Azad Shokrollahi, J. A. Persson, Reza Malekian, Arezoo Sarkheyli-Hägele, and F. Karlsson, "Passive Intfrared Sensor-Based Occupancy Monitoring in Smart Buildings: A Review of Methodologies and Machine Learning Approaches," Sensors, vol. 24, no. 5, pp. 1533–1533, Feb. 2024, doi: https://doi.org/10.3390/s24051533.

[14] GeeksforGeeks, "Supervised vs Reinforcement vs Unsupervised," GeeksforGeeks, Sep. 19, 2024. https://www.geeksforgeeks.org/supervised-vs-reinforcement-vs-unsupervised/

[15] W. Wang, J. Chen, and T. Hong, "Occupancy prediction through machine learning and data fusion of environmental sensing and Wi-Fi sensing in buildings," Automation in Construction, vol. 94, pp. 233–243, Oct. 2018, doi: https://doi.org/10.1016/j.autcon.2018.07.007.

[16] B. Sirmacek and M. Riveiro, "Occupancy Prediction Using Low-Cost and Low-Resolution Heat Sensors for Smart Offices," Sensors, vol. 20, no. 19, p. 5497, Sep. 2020, doi: https://doi.org/10.3390/s20195497.

[17] D. Hong, Q. Gu, and K. Whitehouse, "High-dimensional Time Series Clustering via Cross-Predictability.," International Conference on Artificial Intelligence and Statistics, pp. 642–651, Apr. 2017.

[18] JavaTpoint, "Machine Learning Random Forest Algorithm - Javatpoint," www.javatpoint.com, 2021. https://www.javatpoint.com/machine-learning-random-forest-algorithm