



International Journal of Engineering Research and Science & Technology

www.ijerst.org

ISSN : 2319-5991

Vol. 22 No. 4 (2026)



**ijerst.editor@gmail.com
editor@ijerst.com**

Research Paper

UNIFIED HEALTHCARE DATA INFRASTRUCTURE FOR PATIENT FLOW, BED UTILIZATION, CLINICAL OPERATIONS AND HOSPITAL PERFORMANCE ANALYTICS - (HOSPITAL PULSE)

¹ Mrs. Y. Ramya, ² Ponnaboina Akhila, ³ Gottemukkala Bhavya, ⁴ Peniganbla Yesaswini Chandr, ⁵ Rathod Chandu Naik

¹Assistant Professor, ²³⁴⁵Students

Department of CSE (Data Science)

Siddhartha Institute of Technology & Sciences, Narapally

sits.ramyayelapati@siddhartha.co.in, 24tq1a6738@siddhartha.co.in, 24tq1a6720@siddhartha.co.in,
24tq1a6760@siddhartha.co.in, 24tq1a6739@siddhartha.co.in

ABSTRACT

Hospitals generate a large volume of operational data every day through registrations, admissions, transfers, discharges, laboratory orders, operation theatre schedules, and billing. This data is usually spread across separate systems for the outpatient desk, wards, laboratory, pharmacy, and finance, and each department sees only its own part of the picture. As a result, questions such as how many beds will be free this evening, why emergency patients wait for admission, or which ward has the longest length of stay are answered slowly and often from memory. This paper presents Hospital Pulse, a unified healthcare data infrastructure that integrates these sources and provides analytics for patient flow, bed utilisation, clinical operations, and hospital performance.

The infrastructure is built around an extract, transform, and load pipeline together with change data capture from the hospital information system. Admission, discharge, and transfer events, bed status updates, laboratory and radiology orders, theatre bookings, and billing records are ingested into a raw layer, standardised into a common patient encounter model, and loaded into a clinical operations data warehouse. Patient identifiers are pseudonymised before data enters the analytical layer, and role-based access controls limit what each user can see.

On the integrated data, the system computes key performance indicators such as bed occupancy rate, average length of stay, bed turnover, emergency department waiting time, theatre utilisation, and laboratory turnaround time. Predictive models add a forward-looking view. A gradient boosting model estimates each inpatient's remaining length of stay and the probability of discharge within 24 hours, and a time-series model forecasts daily admissions by department. Together they give a projection of bed availability for the next day, which bed managers can use to plan admissions and transfers. The same data also shows where time is lost during the day, for example between a discharge decision and the release of the bed after billing clearance and cleaning.

I. INTRODUCTION

Efficient patient flow is central to the functioning of a hospital. When a patient arrives at the emergency department, is assessed, admitted to a ward, treated, and discharged, each step depends on the availability of beds, staff, diagnostic services, and operation theatres. A delay at one point, such as a late discharge in the morning or slow laboratory results, causes congestion further upstream, so that emergency patients wait on trolleys and planned surgeries are postponed. Managing this flow requires timely information about the state of the whole hospital, not only individual departments.

Most hospitals already use a hospital information system along with separate systems for the laboratory, radiology, pharmacy, and billing. These systems are designed for transactions, such as registering a patient or recording a test result, rather than for analysis, and their databases are optimised for fast writes on current records rather than for queries that scan months of history. Reports are usually produced at the end of the day or month, and combining data across systems requires manual work by the medical records team. Bed status is often maintained on whiteboards or spreadsheets, and nursing supervisors spend considerable time on phone calls to find out which beds are free and which patients are about to be discharged.

A unified data infrastructure can change this by treating hospital operations data as an integrated asset. By capturing events from all source systems, reconciling them into a common model of patient encounters, and storing them in a warehouse designed for analysis, the hospital can compute consistent indicators in near real time. Machine learning models trained on historical encounters can then predict discharges and admissions, turning static reports into forward-looking guidance for bed management and staffing. Because every indicator is computed from the same curated tables, departments no longer argue over whose figures are correct.

This paper describes Hospital Pulse, a data platform for patient flow, bed utilisation, clinical operations, and performance analytics. It covers data ingestion and integration, privacy protection, the warehouse design, indicator computation, predictive modelling, and the dashboards used by hospital staff. The following sections review related studies, analyse the existing and proposed systems, and present the methodology and architecture.

II. LITERATURE SURVEY

Research on hospital patient flow has examined the causes and consequences of congestion. Studies in operations management describe the hospital as a network of queues in which the emergency department, inpatient wards, intensive care units, and operation theatres are linked, so that bottlenecks in one unit propagate to others. Analyses of emergency department crowding consistently identify access block, the inability to move admitted patients to inpatient beds, as the main cause of long waiting times. Several studies have shown that discharges concentrated in the late afternoon reduce bed availability in the morning when demand from emergency and elective admissions is highest. These findings suggest that improving visibility of discharges and bed status across the hospital can reduce waiting times even without adding beds, provided that information is timely and shared among departments responsible for admissions, nursing, and discharge.

Healthcare data integration and warehousing have been studied extensively. Researchers have described the difficulty of combining data from hospital information systems, laboratory systems, radiology systems, and billing, since each uses its own identifiers, codes, and data structures. Standards such as HL7 messaging and FHIR resources have been proposed to support interoperability, and several studies report experiences of building clinical data warehouses using dimensional models with fact tables for encounters, orders, and bed events. Common challenges include duplicate patient records, inconsistent department codes, and missing timestamps for key events such as transfer and discharge. The literature emphasises data quality checks, master patient indexing, and clear definitions of indicators as prerequisites for reliable analytics, and notes that a warehouse designed with these principles can serve both operational dashboards and retrospective research.

Prediction of length of stay and discharge has been an active area of research. Early studies used regression models based on diagnosis, age, and admission type to estimate length of stay at admission. More recent work has applied random forests, gradient boosting, and neural networks to features collected during the stay, such as vital signs, laboratory results, medication changes, and nursing notes, to predict the probability of discharge within the next day. Studies report that models updated daily with recent clinical information perform better than static admission-time models. Researchers also note that predictions are most useful when presented at ward level as expected discharges, because bed managers plan capacity rather than individual patient outcomes, and that model performance must be checked regularly because clinical practices and patient mix change over time.

Forecasting of hospital admissions and bed occupancy has been approached with time-series and machine learning methods. Seasonal ARIMA, exponential smoothing, and Prophet-style models have been applied to daily emergency attendances and admissions, capturing day-of-week effects and seasonal patterns such as winter respiratory illness and monsoon infections. Some studies have combined admission forecasts with predicted discharges to estimate future occupancy, and simulation models have been used to test the effect of policies such as earlier discharges or additional short-stay beds. The literature shows that combining forecasts at the ward level produces useful estimates of bed pressure one to three days ahead, which supports decisions on elective admissions, staffing, and transfers to other facilities during periods of high demand.

Hospital performance measurement and dashboards form a further research theme. Studies describe indicators including bed occupancy, average length of stay, bed turnover interval, theatre utilisation, laboratory turnaround time, and readmission rates, and discuss how they are used by administrators and accreditation bodies. Research on command centres and operational dashboards in hospitals reports improvements in patient flow when real-time information is displayed to bed managers and department heads, but also warns that dashboards overloaded with indicators are rarely used. Privacy research emphasises pseudonymisation and role-based access when patient data is used for operations analytics. The present work combines these strands by building an integrated, privacy-aware data infrastructure that supports both performance indicators and predictive patient flow analytics within one platform.

III. SYSTEM ANALYSIS

The system is analysed as a data infrastructure that connects transactional hospital systems with operational analytics. The hospital information system produces admission, discharge, and transfer events and holds patient demographic and encounter data. The bed management module records the status of each bed, including occupied, vacant, under cleaning, and blocked. The laboratory and radiology systems provide order and result timestamps, the theatre scheduling system provides bookings and actual start and end times, and the billing system provides charges and discharge clearance times. The ingestion stage captures changes from these systems through change data capture and scheduled extracts. The integration stage matches records to patient encounters using a master patient index, pseudonymises identifiers, and standardises department and ward codes. The warehouse stores encounter, bed event, order, and theatre facts with shared dimensions. Indicator jobs compute performance measures, predictive models estimate discharges and admissions, and the dashboard presents live census, forecasts, and trends to authorised users.

Existing System

In the existing situation, each department in the hospital relies on its own system and reports. The admission desk tracks registrations, the wards maintain bed status on whiteboards or spreadsheets, the laboratory monitors its own turnaround, and the billing department records discharge clearance separately. When the emergency department needs a bed, staff contact ward supervisors by telephone to ask about vacancies and expected discharges. Hospital-wide indicators such as bed occupancy and average length of stay are compiled manually by the medical records department at the end of each month, often from inconsistent sources. Because these reports arrive late, they help in reviewing past performance but not in managing the current day. There is no forecasting of admissions or discharges, so elective admissions and staffing are planned on fixed schedules regardless of expected demand. Data quality problems such as missing discharge times and duplicate registrations remain hidden, and management has little evidence on where delays in patient flow actually occur.

Disadvantages of Existing System

- Operational data is locked in separate departmental systems.
- Bed status is tracked manually on whiteboards and spreadsheets.
- Finding a free bed requires repeated telephone calls between departments.
- Performance indicators are compiled only at the end of the month.
- Admissions and discharges are not forecast in advance.
- Emergency patients wait for beds without visibility of expected discharges.
- Indicator definitions differ between departments and reports.
- Data quality problems such as missing timestamps remain unnoticed.
- Bottlenecks in laboratory and theatre workflows are hard to identify.
- Patient data is shared in reports without systematic privacy controls.

Proposed System

The proposed system, Hospital Pulse, is a unified data infrastructure that integrates operational data from across the hospital and provides real-time and predictive analytics for patient flow. Change data capture streams admission, discharge, transfer, and bed status events from the hospital information system, while scheduled extracts

collect laboratory, radiology, theatre, and billing data. An integration layer matches records to encounters through a master patient index, pseudonymises personal identifiers, and standardises codes. The data is loaded into a clinical operations warehouse built on a dimensional model. Indicator jobs compute occupancy, length of stay, turnover, emergency waiting time, theatre utilisation, and laboratory turnaround every fifteen minutes. A gradient boosting model predicts the probability of each inpatient being discharged within 24 hours, and a time-series model forecasts admissions by department. These are combined into a projected bed availability for each ward. A command centre dashboard and department views present live census, predicted discharges, forecasts, and trends, with access controlled by role.

Advantages of Proposed System

- Brings admission, bed, laboratory, theatre, and billing data into one model.
- Shows live bed status and census for every ward on one screen.
- Predicts likely discharges within the next 24 hours for each ward.
- Forecasts daily admissions by department for capacity planning.
- Projects next-day bed availability to support admission decisions.
- Computes consistent performance indicators every few minutes.
- Highlights delays in laboratory turnaround and discharge clearance.
- Protects patient identity through pseudonymisation and role-based access.
- Exposes data quality issues such as missing timestamps for correction.
- Supports evidence-based review of hospital performance by management.

IV. METHODOLOGY

The methodology begins with a study of source systems to identify tables and events that describe patient movement, bed status, orders, theatre activity, and billing. Change data capture is configured on admission, transfer, discharge, and bed tables, while other sources are extracted every fifteen minutes. Raw records are written to a landing layer and then validated for missing keys, invalid codes, and timestamps that occur in the wrong order, such as a discharge before admission. Valid records are linked to encounters through the master patient index, identifiers are replaced with pseudonymous keys, and department codes are mapped to a standard list. The data is loaded into fact tables for encounters, bed events, orders, and theatre cases with dimensions for date, time, ward, department, and bed. Indicator queries are scheduled to refresh aggregates. For prediction, daily snapshots of each inpatient are built with features such as age group, admission type, department, days since admission, pending tests, and recent procedures, and a LightGBM classifier is trained with time-based validation. Admission forecasts use a seasonal model with holiday effects.

SYSTEM ARCHITECTURE

The architecture of Hospital Pulse is arranged as a hub-and-spoke design around the clinical operations warehouse. The source spokes are the hospital information system, the bed management module, the laboratory and radiology systems, the theatre scheduling system, and billing. Each source connects to the hub through an ingestion channel, using change data capture for bed and movement events and scheduled extracts for the others. At the hub, the integration and privacy layer applies the master patient index, pseudonymisation, and code standardisation before data is stored in the

warehouse. Around the hub, analytical services consume the curated data: the indicator engine for occupancy, length of stay, and turnaround measures, the discharge prediction model, the admission forecasting model, and the bed availability projector. The outputs are served through an API to the command centre dashboard, ward and department views, and scheduled management reports, with role-based access control applied at the API.

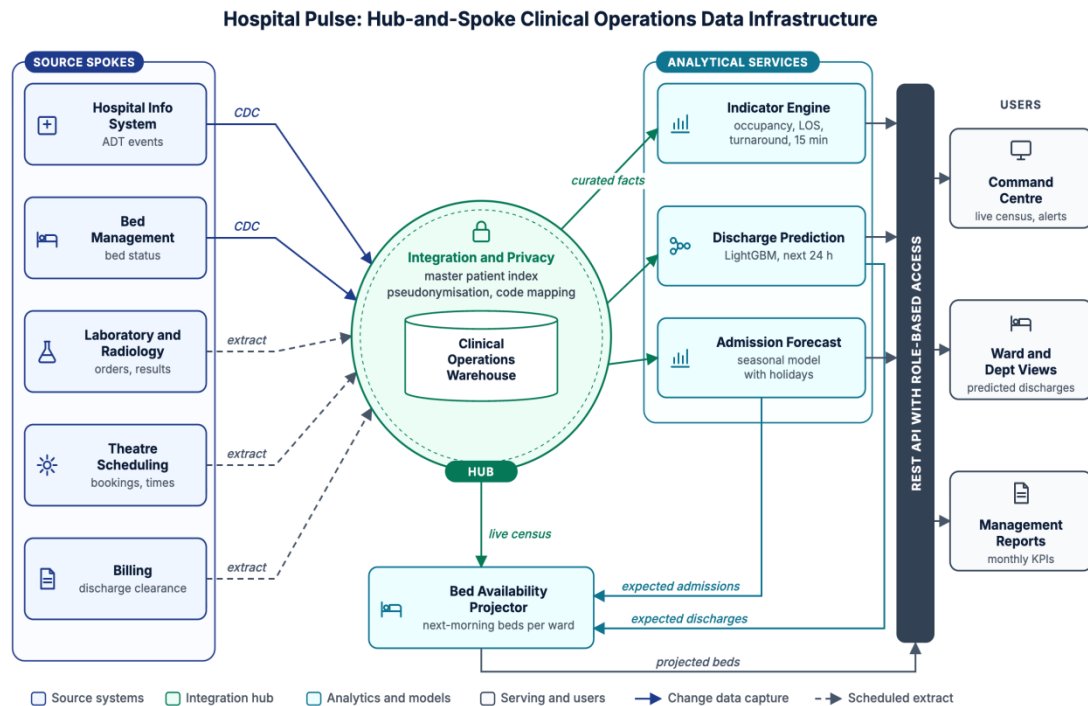


Fig. 1 System Architecture

V. RESULT AND OUTPUT



Fig. 2 Result and Output Screen

VI. CONCLUSION

This paper presented Hospital Pulse, a unified healthcare data infrastructure for patient flow, bed utilisation, clinical operations, and hospital performance analytics. By capturing events from the hospital information system and departmental systems, integrating them into a common encounter model, and storing them in an analytical warehouse, the platform provides a single and consistent view of hospital operations that was previously spread across separate systems, whiteboards, and monthly reports.

The predictive components extend this view into the near future. The discharge prediction model achieved an area under the curve of 0.87, the admission forecast had a mean absolute percentage error of about 7.9 percent, and the combined bed availability projection for the following morning was within three beds of the actual figure on most days of the evaluation period. These results indicate that bed managers can plan admissions and transfers with better information, reducing the time emergency patients spend waiting for inpatient beds.

The indicator engine gives administrators consistent measures that are updated throughout the day, and it has also revealed data quality issues such as missing discharge times that had gone unnoticed in manual reporting. Pseudonymisation and role-based access ensure that operational analytics does not expose patient identity beyond what each role requires. The system supports planning and coordination only, and all clinical decisions, including the decision to discharge a patient, remain with the treating doctors and nursing staff on the ward.

Future work will integrate staff rostering data so that bed forecasts can be matched with nursing capacity, adopt FHIR-based interfaces to simplify connection with other hospitals, and add real-time tracking of equipment such as ventilators and infusion pumps. A multi-site deployment would also allow comparison of performance across

hospitals in a network and support transfers between them during periods of high demand. Such a network view would also help health authorities plan bed capacity for seasonal surges such as monsoon infections and winter respiratory illness.

REFERENCES

- [1] Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In *Handbook of Artificial Intelligence and Wearables* (pp. 145-158). CRC Press.
- [2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In *The International Conference on Artificial Intelligence and Smart Environment* (pp. 557-564). Cham: Springer Nature Switzerland.
- [3] Sv satyakrishna, shirisha rangu, bhargavi nalacheruve. (2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm
- [4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.
- [5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, "Real-Time Object Detection in Drone Surveillance Using YOLOv5," in *Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT)*, Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.
- [6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, "Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks," in *Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment*, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.
- [7] R. D. Kumar, V. N. S. Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in *Blockchain for Smart Cities*, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.
- [8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," *International Research Journal of Modernization in Engineering Technology and Science*, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.
- [9] Ravi Kumar Banoth, Ramana Murthy B V, "Automatic crop recommendation system using LightGBM and decision tree machine learning models," *Journal of Machine and Computing*, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.
- [10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in *Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE)*, Apr. 2025.
- [11] Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," *SN Computer Science*, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.