

ISBN: 978-81-69967-09-9

Artificial Intelligence and Beyond:

Data, Computing and Smart Systems



Editors _____

Dr. Med Ram Verma

Dr. K. Pushpanjali Patra

Dr. Renuka P. Chopde

Er. Harshit Gupta



Bhumi Publishing, India
First Edition: September 2026

Artificial Intelligence and Beyond: Data, Computing and Smart Systems

(ISBN: 978-81-69967-09-9)

DOI: <https://doi.org/10.5281/zenodo.23050973>

Editors

Dr. Med Ram Verma

Principal Scientist and Head,
ICAR-Indian Agricultural Statistics
Research Institute, Pusa, New Delhi

Dr. K. Pushpanjali Patra

Department of Physics,
Ramdeobaba University,
Nagpur, M.S.

Dr. Renuka Prafulla Chopde

Department of Chemistry,
S. Chandra Mahila Mahavidyalaya,
Sakoli, Dist. Bhandara, M.S.

Er. Harshit Gupta

Department of Computer Application,
Rajshree Institute of Management &
Technology, Bareilly, U.P.



Bhumi Publishing

September 2026

Copyright © Editors

Title: Artificial Intelligence and Beyond: Data, Computing and Smart Systems

Editors: Dr. Med Ram Verma, Dr. K. Pushpanjali Patra,

Dr. Renuka Prafulla Chopde, Er. Harshit Gupta

First Edition: September 2026

ISBN: 978-81-69967-09-9



DOI: <https://doi.org/10.5281/zenodo.23050973>

All rights reserved. No part of this publication may be reproduced or transmitted, in any form or by any means, without permission. Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

Published by Bhumi Publishing,

a publishing unit of Bhumi Gramin Vikas Sanstha



Nigave Khalasa, Tal – Karveer, Dist – Kolhapur, Maharashtra, INDIA 416 207

E-mail: bhumipublishing@gmail.com



Disclaimer: The views expressed in the book are of the authors and not necessarily of the publisher and editors. Authors themselves are responsible for any kind of plagiarism found in their chapters and any related issues found with the book.

PREFACE

Artificial Intelligence and Beyond: Data, Computing and Smart Systems presents a broad overview of the rapidly evolving landscape of artificial intelligence, data-driven technologies, advanced computing, and intelligent digital systems. As technological innovation continues to reshape education, industry, healthcare, business, governance, communication, and everyday life, understanding these developments has become increasingly important for researchers, students, professionals, technology enthusiasts, and society at large.

The book brings together diverse perspectives on contemporary developments in artificial intelligence and related fields. It explores emerging approaches in machine learning, data science, intelligent automation, computational technologies, smart systems, digital transformation, and other interdisciplinary applications. By connecting theoretical concepts with practical perspectives, the chapters highlight how data and computing capabilities are enabling increasingly intelligent, adaptive, and connected environments.

A central theme of the book is the growing convergence of artificial intelligence with data, computing infrastructure, and smart technologies. The availability of large and diverse datasets, advances in computational power, and the development of intelligent algorithms are creating new opportunities for innovation and problem-solving. At the same time, these developments raise important considerations related to ethics, privacy, security, reliability, accessibility, and responsible technological use.

The chapters included in this volume reflect ongoing research, emerging trends, practical applications, and future possibilities across different domains. The interdisciplinary nature of the contributions encourages readers to view artificial intelligence not as an isolated technology, but as part of a wider technological ecosystem.

It is hoped that this book will serve as a useful academic and reference resource for students, researchers, educators, professionals, and practitioners. By presenting current ideas and diverse applications, the volume seeks to encourage further research, collaboration, innovation, and responsible development of intelligent technologies for a digitally connected future today.

- Editors

TABLE OF CONTENT

Sr. No.	Book Chapter and Author(s)	Page No.
1.	AI-DRIVEN ADAPTIVE MULTILAYER SECURITY ARCHITECTURE FOR PRIVACY-PRESERVING THREAT DETECTION IN CLOUD-BASED BIG DATA ANALYTICS K. Srinath, Siva Kumar Dhandapani and R. Sundar	1 – 11
2.	ARTIFICIAL INTELLIGENCE FOR GENETIC IMPROVEMENT OF LIVESTOCK Anil Chitra, Yogesh C. Bangar, Kamaldeep and Rakesh Nehra	12 – 23
3.	AI-BASED HEALTH MONITORING OF FARM ANIMALS Sunil Kumar, Yogesh C. Bangar, Anil Chitra and Nitesh Jat	24 – 31
4.	APPLICATION SECURITY CHALLENGES IN AI-ENABLED SOFTWARE SYSTEMS: RISKS, CONTROLS, AND SECURE DEVELOPMENT PRACTICES Udaya Bhaskar Vemuri	32 – 40
5.	ARTIFICIAL INTELLIGENCE IN VETERINARY DIAGNOSTIC IMAGING: CURRENT APPLICATIONS, CHALLENGES AND FUTURE PERSPECTIVES Antra Ken, Sangeet Rana and Aashish	41 – 52
6.	TALENT ACQUISITION IN THE ERA OF ARTIFICIAL INTELLIGENCE: A REVIEW OF CURRENT TRENDS AND CHALLENGES Monika Chauhan and Nisha Agarwal	53 – 63
7.	AI-DRIVEN SMART SYSTEMS IN BUSINESS MANAGEMENT Divyansh Mishra, Rajesh Kumar Mishra and Rekha Agarwal	64 – 75
8.	STATISTICAL METHODS FOR COMPUTING AND DATA SCIENCE: FOUNDATIONS, ALGORITHMS AND APPLICATIONS Umesh R. Yeole	76 – 89
9.	DATA ANALYTICS FOR ANIMAL HEALTH: APPLICATIONS, CHALLENGES AND FUTURE PERSPECTIVES Yogesh C. Bangar, S. S. Dhaka, Sunil Kumar and Yash Raj	90 – 97

10.	ARTIFICIAL INTELLIGENCE AND ORGANIZATIONAL TRANSFORMATION: A HUMANISTIC APPROACH TO CONTEMPORARY MANAGEMENT	98 – 105
	Ranjit Kumar Paswan	
11.	AI-DRIVEN HUMAN RESOURCE MANAGEMENT AND EMPLOYEE WELL-BEING: AN INTERDISCIPLINARY PERSPECTIVE ON THE FUTURE OF WORK	106 – 118
	V. T. Vidhya Poorani and S. Subramanian	
12.	IOT-BASED GREENHOUSE CONTROLLING AND MONITORING SYSTEM USING ARDUINO UNO AND ESP8266	119 – 134
	Shalu C., Mohammad Altamas and Debika Chaudhuri	
13.	DESIGN AND IMPLEMENTATION OF A PRIVACY-PRESERVING IDENTITY MECHANISM FOR FILE STORAGE SERVERS	135 – 141
	Devayani Shailendra Singh Girase, Vaishnavi Sandip Jagtap, Hiraman Jadhav, Mayuri Rathi and Ruchi Thakare	

AI-DRIVEN ADAPTIVE MULTILAYER SECURITY ARCHITECTURE FOR PRIVACY-PRESERVING THREAT DETECTION IN CLOUD-BASED BIG DATA ANALYTICS

K. Srinath¹, Siva Kumar Dhandapani² and R. Sundar³

¹Department of CSE, AMET University, Chennai, India 603112

²Department of CSE, Cyber Security,

Rajalakshmi Engineering College, Chennai, India 602105

³Department of Marine Engineering, AMET University, Chennai, India 603112

Corresponding author E-mail: srinath8031@ametuniv.ac.in,
sivakumar.d@rajalakshmi.edu.in, sundar.r@ametuniv.ac.in

Abstract

Cloud-based big data analytics has become a central component of modern digital services because it provides elastic computing, distributed storage, and rapid processing of large and heterogeneous datasets. At the same time, the combination of cloud infrastructure, big data, and intelligent analytics creates a broad attack surface. Sensitive information may be exposed through compromised accounts, insecure application programming interfaces, misconfigured storage, malicious insiders, data poisoning, and advanced network attacks. Conventional security mechanisms based mainly on static rules and signatures are often inadequate for detecting changing and previously unseen threats. This chapter presents an AI-driven adaptive multilayer security architecture for privacy-preserving threat detection in cloud-based big data analytics. The architecture combines data-level protection, secure communication, cloud infrastructure security, big data security, artificial intelligence-based detection, and adaptive response. Machine learning and deep learning methods are considered for anomaly and intrusion detection, while encryption, anonymization, differential privacy, and federated learning are used to reduce privacy exposure. A risk-aware adaptive controller continuously evaluates threat evidence and adjusts security controls according to the observed context. The chapter describes the architecture, threat model, detection workflow, privacy mechanisms, algorithm, evaluation metrics, implementation considerations, limitations, and future research directions. The proposed design provides a practical foundation for developing secure, scalable, intelligent, and privacy-aware cloud analytics environments.

Keywords: Artificial Intelligence; Big Data Analytics; Cloud Computing; Cybersecurity; Intrusion Detection; Privacy Preservation; Machine Learning; Deep Learning; Adaptive Security; Federated Learning.

1. Introduction

The volume and variety of digital information generated by organizations, connected devices, online services, and enterprise applications have increased substantially. Big data analytics enables organizations to transform this information into useful knowledge by applying distributed storage, parallel processing, statistical analysis, and machine learning. Cloud computing complements these capabilities by providing on-demand computing resources and flexible storage without requiring organizations to maintain all physical infrastructure locally. As a result, cloud-based analytics is increasingly used in healthcare, banking, education, manufacturing, transportation, e-commerce, and smart-city applications.

The advantages of cloud-based big data analytics are accompanied by significant security and privacy concerns. Data may move between users, applications, cloud services, storage systems, and analytical engines. A weakness in any component can affect the confidentiality, integrity, or availability of the overall system. Common threats include credential theft, unauthorized access, malware, denial-of-service attacks, data leakage, insider misuse, insecure APIs, privilege escalation, and exploitation of cloud misconfigurations. In analytical environments, attackers may additionally target the datasets and machine learning models themselves through poisoning, evasion, or inference attacks.

Artificial intelligence offers a promising approach for addressing these challenges. Unlike purely signature-based systems, AI models can learn statistical relationships and behavioral patterns from security data. Classification models can distinguish known malicious and benign events, while anomaly-detection models can identify deviations from normal behavior. However, AI-based security must itself be protected. Training data can contain sensitive information, models can be manipulated, and centralized collection of security logs can create additional privacy risks.

This chapter therefore proposes a multilayer architecture in which AI is treated as one component of a broader defense-in-depth strategy. Privacy protection is incorporated before and during analytics, and an adaptive security controller uses threat scores and contextual information to select appropriate responses. The goal is not to replace established security controls but to coordinate them with intelligent detection so that the system can respond to changing conditions while limiting unnecessary disclosure of sensitive data.

2. Background and Related Concepts

2.1 Cloud-Based Big Data Analytics

Cloud-based big data analytics combines distributed data storage and processing with elastic cloud resources. Large datasets can be divided across multiple nodes, processed in parallel, and consolidated into analytical results. This architecture improves scalability but increases the number of services, identities, interfaces, and communication paths that require protection.

2.2 Artificial Intelligence for Cybersecurity

AI-based cybersecurity uses computational models to recognize patterns associated with attacks or abnormal behavior. Supervised learning is useful when representative labeled datasets are available. Unsupervised and semi-supervised approaches are useful when attack labels are incomplete. Deep learning can model complex relationships in high-dimensional or sequential security data. The appropriate model depends on the data characteristics, computational budget, interpretability requirements, and operational context.

2.3 Privacy-Preserving Analytics

Privacy-preserving analytics aims to obtain useful analytical results without unnecessarily exposing sensitive raw information. Encryption, access control, data masking, anonymization, differential privacy, and federated learning provide different levels of protection. No single technique is sufficient for every application; therefore, privacy controls should be selected according to data sensitivity, analytical requirements, and threat exposure.

3. Security and Privacy Challenges

- Data confidentiality: sensitive records may be exposed during storage, transmission, processing, or sharing.
- Data integrity: manipulation of datasets, logs, or analytical results can produce incorrect decisions.
- Identity and access risks: stolen credentials and excessive privileges can enable unauthorized activity.
- Insider threats: legitimate accounts may be misused to access or exfiltrate sensitive information.
- Distributed attack surface: cloud services, containers, virtual machines, APIs, and storage systems create many possible entry points.
- Advanced attacks: attackers can use automated, evasive, and previously unseen techniques that are difficult to detect with fixed signatures.
- Training-data poisoning: manipulated training examples can reduce the reliability of AI-based detectors.
- Model evasion: carefully crafted inputs may cause a model to classify malicious activity as legitimate.
- Privacy leakage: centralized security analytics may expose personally identifiable or business-sensitive information.
- Concept drift: normal behavior and attack patterns change over time, reducing the effectiveness of static models.

4. Proposed Adaptive Multilayer Security Architecture

The proposed architecture follows a defense-in-depth model with coordinated controls across eight logical layers: data acquisition, privacy preservation, secure communication, cloud infrastructure, big data analytics, AI threat detection, adaptive response, and security management. The layers are connected through a policy and risk engine. Security events generated at lower layers are transformed into privacy-protected features and supplied to the AI detection layer. The resulting threat score is combined with contextual information and used to select a response.

4.1 Data Acquisition Layer

The data acquisition layer receives information from databases, applications, IoT devices, network sensors, authentication services, cloud logs, and security tools. Before analysis, data is classified according to sensitivity and relevance. Examples of security features include failed-login counts, access frequency, session duration, API request patterns, data-transfer volume, unusual geographic activity, and resource-utilization behavior.

4.2 Privacy Preservation Layer

This layer applies protection according to the sensitivity of the information. Encryption protects stored and transmitted data. Masking and pseudonymization can reduce direct exposure of identifiers. Differential privacy can limit disclosure from analytical outputs. Federated learning can allow multiple sites to train a shared detector while keeping raw training data at the source. Privacy controls should be applied selectively because stronger protection can increase computational and communication overhead.

4.3 Secure Communication Layer

Secure communication is required whenever information moves between data sources, cloud services, analytical nodes, and security components. Mutual authentication, certificate management, encrypted transport, integrity checking, secure APIs, and network segmentation can reduce interception and impersonation risks. A zero-trust approach can be used so that access is evaluated continuously rather than being granted solely because a user or device is inside a trusted network boundary.

4.4 Cloud Infrastructure Security Layer

The infrastructure layer protects virtual machines, containers, storage, networks, identities, and cloud management interfaces. Role-based or attribute-based access control should follow least-privilege principles. Configuration assessment, vulnerability management, network segmentation, logging, and resource isolation provide additional protection. Automated configuration checks can help identify exposed storage, excessive permissions, or insecure services before they are exploited.

4.5 Big Data Analytics Security Layer

The analytics layer protects distributed processing pipelines and analytical datasets. Important controls include dataset authorization, secure ingestion, integrity verification, query auditing, data lineage, and separation of analytical workloads. Data lineage is useful for investigating suspicious changes because it records the origin and transformation history of analytical information.

4.6 AI-Based Threat Detection Layer

The AI layer processes privacy-protected security features and estimates the probability or severity of malicious behavior. Supervised models such as Random Forest, Support Vector Machine, and gradient-boosting classifiers can be used for labeled attack categories. Autoencoders and clustering can support anomaly detection when labels are limited. Recurrent or transformer-based models can be considered for sequential event streams. Model selection should be based on empirical evaluation rather than assuming that a more complex model is always better.

4.7 Adaptive Response Layer

The adaptive layer translates threat predictions into security actions. A low-risk event may simply be logged. A moderate-risk event may trigger enhanced monitoring or step-up authentication. A high-risk event can result in temporary access restriction, session termination, or network isolation. Critical events may trigger resource quarantine and incident-response procedures. The controller should consider confidence, asset criticality, user context, and historical evidence so that an unusual but legitimate activity is not automatically treated as a confirmed attack.

4.8 Security Management and Monitoring Layer

The management layer provides centralized visibility into alerts, model performance, policy changes, privacy controls, and response actions. It maintains audit records and supports human review. Security analysts should be able to override automated actions when the consequences of an incorrect response are significant.

5. Threat Model

The architecture assumes an attacker may target any exposed component of the cloud analytics environment. External attackers may attempt credential attacks, malicious API requests, malware delivery, denial of service, or data exfiltration. Internal attackers may misuse legitimate privileges. An attacker may also attempt to poison AI training data or generate inputs designed to evade detection. The model therefore considers confidentiality, integrity, availability, identity, privacy, and AI-model security as related objectives.

The architecture does not assume that every abnormal event is malicious. Legitimate workloads can produce unusual patterns during software releases, disaster recovery, business peaks, or

maintenance. For this reason, the adaptive controller incorporates contextual information and confidence thresholds rather than relying on a single anomaly score.

6. AI-Based Intelligent Threat Detection

6.1 Data Preprocessing

Security data often contains missing values, duplicated events, inconsistent formats, and highly imbalanced attack classes. Preprocessing may include deduplication, missing-value handling, normalization, categorical encoding, temporal aggregation, and removal of irrelevant fields. Privacy-sensitive identifiers should be transformed before model training whenever they are not required for the analytical objective.

6.2 Feature Engineering

Useful features can describe frequency, timing, sequence, volume, and relationships between events. Examples include the number of failed authentications within a time window, average data-transfer size, unusual API-call sequences, number of accessed resources, and deviation from a user's historical behavior. Feature selection is important because unnecessary features increase computational cost and may introduce privacy exposure.

6.3 Detection Models

A hybrid strategy can combine supervised classification with anomaly detection. The classifier handles patterns represented in labeled training data, while an anomaly detector provides a mechanism for identifying behavior that differs from established normal patterns. An ensemble can improve robustness when individual models make different types of errors. Model calibration is important because an output probability should not automatically be interpreted as a reliable probability of attack without validation.

6.4 Threat Scoring

A simplified threat score can be expressed as:

$$T = w_1A + w_2B + w_3C + w_4D$$

where A represents the AI prediction, B represents behavioral deviation, C represents asset criticality, D represents contextual risk, and the weights w_1 – w_4 represent the relative importance of these factors. In a practical implementation, the score can be normalized to a common range and mapped to response levels.

6.5 Continuous Learning

Cloud environments evolve continuously. New applications, users, devices, and attack techniques can change the underlying data distribution. The architecture should therefore monitor model drift and periodically retrain or update models using validated events. Automatic learning from every alert should be avoided because poisoned or incorrectly labeled events can degrade the detector.

7. Privacy-Preserving Security Mechanisms

7.1 Encryption

Encryption protects confidentiality when information is stored or transmitted. Key management is as important as the encryption algorithm itself. Keys should be protected separately from the encrypted data, rotated according to policy, and access-controlled.

7.2 Anonymization and Pseudonymization

Anonymization attempts to prevent information from being associated with an individual, while pseudonymization replaces direct identifiers with artificial identifiers. These techniques are useful for analytics but should not be treated as universal protection against re-identification, especially when multiple datasets can be linked.

7.3 Differential Privacy

Differential privacy provides a mathematical framework for limiting the influence of an individual record on an output. A simplified mechanism is $M(D) = f(D) + N$, where $f(D)$ is the desired analytical function and N is calibrated noise. Privacy strength and analytical utility must be balanced carefully.

7.4 Federated Learning

Federated learning distributes model training across participating data holders. Each participant trains a local model and sends selected model updates to an aggregation process. A basic weighted aggregation can be represented as $W = \sum(m_i/M)W_i$. Federated learning reduces the need to centralize raw data, although model updates can still leak information unless additional protections such as secure aggregation or differential privacy are applied.

7.5 Access Control

Privacy is also supported through authorization. Data should be accessible only to identities and services with a legitimate requirement. Attribute-based policies can incorporate user role, data classification, device status, location context, and purpose of access. Fine-grained authorization is particularly important in multi-tenant cloud analytics.

8. Adaptive Threat Response

The response engine uses threat severity and contextual risk to select controls. An example policy is shown below.

Threat Level	Example Condition	Adaptive Response
Low	Minor deviation	Log and continue monitoring
Moderate	Repeated abnormal behavior	Increase monitoring and request additional authentication
High	Strong evidence of compromise	Restrict session, revoke selected privileges, alert analyst
Critical	Confirmed or highly probable attack	Isolate resource, block activity, initiate incident response

Adaptive response should be reversible where possible. Temporary restrictions reduce the operational impact of false positives, while high-impact actions should require additional evidence or human approval. The response engine should record the reason for every automated action to support auditing and model improvement.

9. Proposed Algorithm

Algorithm 1: AI-Driven Adaptive Privacy-Preserving Threat Detection

Input: Cloud security events X

Output: Threat category and security response

- i. Collect events from cloud, network, application, identity, and analytics sources.
- ii. Classify each event and associated data according to sensitivity.
- iii. Apply encryption, masking, anonymization, differential privacy, or distributed learning as appropriate.
- iv. Preprocess the protected security data.
- v. Extract behavioral and contextual features.
- vi. Apply the trained AI detection model.
- vii. Calculate a combined threat score using model output and contextual risk.
- viii. Check model confidence and asset criticality.
- ix. Assign the event to a low, moderate, high, or critical risk category.
- x. Select the least disruptive response that provides adequate protection.
- xi. Apply the response and record the event in an immutable or access-controlled audit trail.
- xii. Send validated outcomes to the model-monitoring component.
- xiii. Detect model drift and initiate controlled retraining when required.
- xiv. Repeat continuously.

10. Performance Evaluation

A practical implementation should evaluate both cybersecurity effectiveness and privacy cost. Accuracy alone is not sufficient because security datasets are frequently imbalanced. The following metrics provide a more complete assessment.

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN)$$

$$\text{Precision} = TP / (TP + FP)$$

$$\text{Recall} = TP / (TP + FN)$$

$$\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

$$\text{False Positive Rate} = FP / (FP + TN)$$

Detection latency should also be measured because a highly accurate detector that responds too slowly may have limited operational value. Resource overhead should be measured in terms of CPU, memory, storage, and network consumption. For privacy-preserving mechanisms,

additional measurements should include communication overhead, encryption cost, privacy budget where applicable, and reduction in raw-data exposure.

A representative experimental design can compare a conventional single-layer detector with the proposed multilayer architecture. The same dataset, preprocessing procedure, and evaluation split should be used for fair comparison. Evaluation should include known attacks and, where possible, previously unseen attack patterns. Results should be reported with confidence intervals or repeated trials rather than relying on a single run.

11. Implementation Considerations

The architecture can be implemented using a combination of cloud-native security services, distributed data-processing platforms, machine learning frameworks, and security information and event management tools. Data ingestion should be separated from model training where possible so that an attack on an analytical workload does not directly compromise the security-management plane.

Containerized AI services can provide portability, while edge or local processing can reduce the transfer of sensitive raw data. A model registry should track model versions, training datasets, evaluation results, and deployment status. Access to model-management functions should be tightly controlled because unauthorized model replacement could undermine the complete security system.

Operational deployment should begin with monitoring-only mode. During this phase, the system generates predictions without automatically blocking users or resources. After the detection quality and false-positive behavior have been validated, selected low-risk automated responses can be enabled. High-impact responses should remain subject to strong policy controls and, where appropriate, human confirmation.

12. Advantages and Limitations

12.1 Advantages

- Defense in depth reduces dependence on a single security control.
- AI enables behavioral and anomaly-based threat detection.
- Adaptive policies can respond differently to different levels of risk.
- Privacy mechanisms reduce unnecessary exposure of sensitive data.
- The architecture can scale with cloud-based analytics workloads.
- Continuous monitoring supports earlier identification of suspicious activity.

12.2 Limitations

- AI models require representative and trustworthy training data.
- False positives can disrupt legitimate cloud workloads.
- Privacy mechanisms may introduce computational or communication overhead.

- Adaptive systems can be difficult to validate because policies change with context.
- AI models remain vulnerable to poisoning, evasion, and other model-specific attacks.
- Integration across heterogeneous cloud platforms can increase implementation complexity.

13. Future Research Directions

Future work can extend the architecture in several directions. Explainable AI can provide analysts with understandable reasons for threat classifications. Federated threat intelligence can enable organizations to learn from distributed security experiences without sharing raw logs. Adversarially robust training can improve resistance to deliberate model manipulation. Edge-cloud cooperation can move initial detection closer to sensitive data sources while retaining cloud-scale analytical capabilities.

Another promising direction is autonomous security orchestration. An AI system could combine threat intelligence, asset criticality, historical incidents, and policy constraints to recommend or execute response actions. Such autonomy should be accompanied by safeguards, auditability, rollback mechanisms, and human oversight. Research into post-quantum cryptography may also become important for long-lived cloud data and security infrastructure.

Future empirical studies should evaluate the architecture on multiple datasets and realistic cloud workloads. Experiments should measure the trade-offs among detection accuracy, privacy, latency, scalability, and operational cost. This multidimensional evaluation is necessary because improving one security property may adversely affect another.

Conclusion

Cloud-based big data analytics provides substantial benefits for large-scale data processing but introduces a complex combination of cybersecurity and privacy challenges. Static security mechanisms alone are unlikely to provide sufficient protection against evolving threats, particularly when cloud services, distributed analytics, and AI systems are tightly interconnected. This chapter presented an AI-driven adaptive multilayer security architecture that integrates privacy preservation with intelligent threat detection. The architecture distributes security controls across data, privacy, communication, cloud infrastructure, analytics, AI detection, adaptive response, and monitoring layers. Machine learning and deep learning can identify suspicious behavior, while encryption, anonymization, differential privacy, and federated learning can reduce unnecessary exposure of sensitive information. The adaptive response layer uses threat severity and contextual information to select appropriate controls and support continuous protection.

The proposed architecture should be viewed as a framework rather than a single algorithm. Its effectiveness depends on careful model selection, trustworthy training data, secure implementation, privacy-aware data handling, and continuous evaluation. By combining AI,

multilayer defense, adaptive response, and privacy-preserving analytics, organizations can move toward cloud environments that are not only intelligent and scalable but also more secure, resilient, and trustworthy.

References

1. Breiman, L. (2001). Random forests. *Machine Learning*, 45, 5–32.
2. Dwork, C. (2006). Differential privacy. In *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming*.
3. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
4. McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. *Proceedings of AISTATS*.
5. Mell, P., & Grance, T. (2011). *The NIST definition of cloud computing*. NIST Special Publication 800-145.
6. National Institute of Standards and Technology. (2023). *Artificial Intelligence Risk Management Framework (AI RMF 1.0)*. NIST AI 100-1.
7. National Institute of Standards and Technology. (2024). *Cybersecurity Framework (CSF) 2.0*. NIST.
8. Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *IEEE Symposium on Security and Privacy*.
9. Stallings, W. (2018). *Effective cybersecurity: A guide to using best practices and standards*. Pearson.
10. Zhou, Y., Cheng, G., Jiang, S., & Dai, M. (2021). Building an efficient intrusion detection system based on feature selection and ensemble classifier. *Computer Networks*.

ARTIFICIAL INTELLIGENCE FOR GENETIC IMPROVEMENT OF LIVESTOCK

Anil Chitra*, Yogesh C. Bangar, Kamaldeep and Rakesh Nehra

Department of Animal Genetics and Breeding,

Lala Lajpat Rai University of Veterinary and Animal Sciences (LUVAS), Hisar 125004

*Corresponding author E-mail: anil.15chitra@luvas.edu.in

Abstract

Artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL), is increasingly being used in livestock genetic improvement as breeding programmes generate growing volumes of genetic, phenotypic, sensor-derived, environmental, and multi-omics data. This chapter examines the role of AI in the evolution of livestock breeding from phenotypic and pedigree-based selection to quantitative genetic evaluation, marker-assisted selection, and genomic selection. AI-enabled high-throughput phenotyping uses cameras and sensors to collect repeated measurements of traits that are difficult, expensive, or subjective to record. Studies of milk production, growth, disease resistance, and fertility indicate that AI-based methods can provide useful predictions, although their performance varies across traits and populations and may not consistently outperform established approaches such as GBLUP. Integrating genetic, phenotypic, environmental, and multi-omics data may provide opportunities to better characterise complex genotype–phenotype relationships and improve genomic prediction, selection, and mating decisions. However, challenges including large and complex datasets, small sample sizes, poor data quality, overfitting, population differences, limited interpretability, high computational requirements, and insufficient validation using independent data remain important. Overall, AI is best viewed as a complementary component of modern livestock breeding that can strengthen phenotyping, genomic prediction, and decision-making when integrated with quantitative genetics and validated across independent populations, generations, and environments. Its ultimate value will depend on whether it contributes to improved selection efficiency, realised genetic gain, and sustainable improvement of production, health, fertility, efficiency, longevity, and environmental resilience.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Livestock Breeding, Genomic Selection, High-Throughput Phenotyping, Genomic Prediction, Precision Livestock Farming.

1. Introduction

Genetic improvement of livestock depends on identifying and selecting animals with superior genetic merit (Garrick & Golden, 2009; Weigel, 2017). Animal breeding has progressively

evolved from phenotypic selection to genetic evaluation, molecular marker-assisted selection and genomic selection, resulting in more accurate estimation of genetic potential (Meuwissen *et al.*, 2001; VanRaden, 2008; Panigrahi *et al.*, 2025).

Artificial intelligence (AI), especially machine learning and deep learning, can help us to analyse large and complex datasets to identify useful information and patterns (Nayak *et al.*, 2024; Pedrosa *et al.*, 2024). One of the immediate applications in animals is high-throughput phenotyping, in which images, video and sensor-based measurements can be converted into phenotypes for traits that are costly or difficult to record manually (Ventura *et al.*, 2020). Machine learning methods have been reported to identify complex patterns effectively and to improve the accuracy of genomic estimated breeding value (GEBV) predictions in some applications (Chafai *et al.*, 2023). Beskorovajni *et al.* (2021) developed an artificial neural network model for predicting yield and fertility traits in Serbian Holstein cattle and reported good predictive performance, although the study was based on a relatively small dataset. Genomic prediction has been investigated using a range of statistical and machine-learning approaches across livestock species, including dairy cattle (Beskorovajni *et al.*, 2021), beef cattle (Srivastava *et al.*, 2021), and pigs (Zhao *et al.*, 2020), with the relative performance of different methods varying across traits and populations. Recent studies have applied machine learning methods to the genomic prediction of complex behavioral traits in dairy cattle, including traits measured using automatic milking systems in North American Holstein cattle (Pedrosa *et al.*, 2024). However, computational requirements, dependence on genomic information, and the need for validation across independent populations remain important challenges for the broader adoption of machine learning in routine breeding programmes (Chafai *et al.*, 2023; Morota *et al.*, 2026).

Thus, these approaches are being used as an alternative in animal breeding and can work together with traditional breeding methods to improve the prediction of economically important traits (Nayeri *et al.*, 2019). Therefore, AI should be regarded as a computational component of modern genetic improvement rather than a complete replacement for quantitative genetics or genomic selection.

2. Evolution of Livestock Genetic Improvement Toward Artificial Intelligence

Early livestock improvement was done by selection based on phenotype and pedigree records (Boichard *et al.*, 2016; Panigrahi *et al.*, 2025). Because phenotypic performance is influenced by both genetic and environmental factors, quantitative genetics uses parameters such as heritability, genetic covariance and breeding values to describe genetic variation, while BLUP incorporates pedigree and phenotypic information to predict breeding values (Zhang *et al.*, 2026; Silva *et al.*, 2026).

The development of molecular markers further provided information on genetic variation. Marker-assisted selection (MAS) incorporated markers associated with quantitative trait loci (QTL); however, most economically important traits are polygenic and exhibit continuous phenotypic variation, making MAS based on individual QTL-associated markers less effective for capturing the overall genetic variation underlying these traits (Panigrahi *et al.*, 2025). This limitation contributed to the development of genomic selection (GS), in which dense genome-wide markers are used to predict total genetic value without requiring prior identification of individual causal loci (Meuwissen *et al.*, 2001).

In the post-genomic era, animal breeding increasingly involves high-dimensional data generated from genomics, epigenomics, transcriptomics, proteomics, and metabolomics (Wadood *et al.*, 2025). Modern breeding programmes also generate genomic, pedigree, production, health, reproductive, and environmental data, while precision-livestock technologies provide additional information through cameras, automated milking systems, and wearable sensors (Morota *et al.*, 2026; Iqbal, 2026). AI can facilitate the integration of these heterogeneous data sources, particularly for complex traits influenced by multiple genetic and environmental factors. Multi-omics approaches can further link genetic variation with molecular processes and phenotypic expression, potentially improving biological understanding and prediction of complex traits (Wadood *et al.*, 2025). However, the high dimensionality of multi-omics data relative to often limited sample sizes in livestock studies can increase the risk of overfitting and biased predictions, making appropriate feature selection, dimensionality reduction, regularisation, and independent validation essential (Rohart *et al.*, 2023).

To analyse such high-dimensional datasets, bioinformatics tools and machine learning methods, including random forests, support-vector machines, kernel methods and neural networks can be used (Tarjan *et al.*, 2021). Deep learning has been particularly useful for image-based livestock phenotyping because convolutional neural networks (CNNs) can automatically learn hierarchical feature representations from image data, reducing the need for manual feature extraction (Montes *et al.*, 2026). Therefore, AI represents an extension of existing breeding systems rather than a replacement.

3. Role of AI in Livestock Genetic Improvement

AI contributes to livestock genetic improvement through high-throughput phenotyping, integration of genomic, phenotypic, and environmental information, genomic prediction, and improved selection decisions. Its application is particularly valuable for traits that are difficult, expensive, labour-intensive, or subjective to measure, such as feed efficiency, fertility, health, and behavioural traits.

One of the most important applications of machine learning in animal breeding is genomic prediction (Chafai *et al.*, 2023). Traditional genetic evaluation methods such as Best Linear

Unbiased Prediction (BLUP) have been widely used to estimate breeding values, while genomic approaches such as Genomic Best Linear Unbiased Prediction (GBLUP) incorporate genome-wide marker information. Machine learning methods, including Random Forests (RF), Support Vector Machines (SVM), and kernel ridge regression (KRR), can model complex or nonlinear relationships and may provide competitive prediction accuracy for some traits and populations (Wang *et al.*, 2022). However, the value of AI in livestock genetic improvement should be assessed not only by predictive accuracy but also by its ability to generate reliable and genetically informative phenotypes, integrate diverse data sources, and improve selection decisions.

4. AI-Enabled High-Throughput Phenotyping

4.1 Digital Phenotyping

Traditional livestock phenotyping can be labour-intensive, subjective, time-consuming, or expensive. Digital phenotyping uses imaging and sensor technologies to obtain objective, quantitative, and repeatable measurements from animals (Neethirajan & Kemp, 2021). Camera-based systems can characterize body shape, posture, and movement; three-dimensional imaging systems can quantify body dimensions and conformation; and wearable or stationary sensors can monitor activity, behaviour, and other physiological or behavioural indicators. Bewley *et al.* (2008) demonstrated the potential of digital images for estimating body condition score in dairy cattle, while Spoliansky *et al.* (2016) developed an automated body-condition scoring system using a low-cost three-dimensional Kinect camera.

4.2 Computer Vision and Sensor-Based Phenotyping

Convolutional neural networks (CNNs) can perform regression, image classification, detection, and segmentation, making them suitable for automated phenotyping (Montes *et al.*, 2026). Their applications include body condition, conformation, locomotion and behavioural traits. High-throughput phenotyping is therefore one of the most immediate contributions of AI to livestock improvement because it increases the quantity, frequency and efficiency of phenotypic data collection for genetic evaluation.

5. AI for Genomic Prediction and Estimation of Genetic Merit

Genomic prediction uses genome-wide marker information to estimate genetic merit. Machine-learning methods such as random forest, support-vector machines, kernel methods and neural networks have been studied because they can handle high-dimensional data and can identify nonlinear relationships (Wang *et al.*, 2022).

However, the performance of AI-based genomic prediction is trait- and population-dependent. Hay *et al.* (2024) compared machine-learning approaches with GBLUP for growth traits in beef cattle and found that GBLUP performed best for birth and yearling weight, whereas random forest performed best for weaning weight. This indicates that AI methods cannot be expected to

consistently outperform established genomic prediction methods. The ability of some AI-based methods to model complex nonlinear relationships makes them potentially useful for traits in which non-additive genetic effects, such as dominance and epistasis, contribute to phenotypic variation (Zhang *et al.*, 2026). Lee *et al.* (2023) developed deepGBLUP, which integrates deep learning with GBLUP to account for complex genetic relationships, and reported competitive or improved predictive performance for several traits in Korean native cattle.

6. AI for Genotype-Phenotype Relationships

Economic traits such as growth, milk production, fertility, disease resistance and feed efficiency are quantitative and are influenced by multiple genes, affected by environmental effects and non-additive genetic interactions.

Methods like Random forests, nonlinear kernels and neural networks can identify interactions without any need to specify all relationships in advance. Epistasis may be of interest because the effect of one locus may depend on another. However, the number of possible interactions increases quickly with number of markers making reliable estimation difficult in small livestock populations.

The main challenge is therefore not simply identifying statistical associations but determining whether they are reproducible and biologically relevant. A genetic marker identified as important by an AI model may predict a trait because of linkage disequilibrium, population structure, or its association with a causal variant. Therefore, a marker showing strong predictive importance should not automatically be interpreted as a causal variant underlying the trait.

AI outputs can be of three different types:

- i. **Pattern recognition** - finding relationships or patterns in the data (e.g.: relationship between body shape measurements and milk yield in dairy cattle.).
- ii. **Phenotypic prediction** - predicting an observed trait or biological condition (e.g.: predicting milk yield of a dairy cow using available information).
- iii. **Genetic prediction** - estimating the inherited genetic potential of an animal for breeding and selection (e.g.: predicting genetic merit of dairy bulls for milk yield using their genetic marker data and dam performance records).

Genotype $\times$ environment interaction (G $\times$ E) provides another important application of AI in livestock breeding. Genetic differences in animal performance may be expressed differently under specific climatic or management conditions. AI can integrate genomic information with environmental and longitudinal phenotypic data, such as repeated milk-yield records, to investigate how animals respond to different environmental conditions. Such applications are relevant to the genetic improvement of resilience and climate adaptation; however, appropriate genetic and statistical analyses are still required to distinguish genetic effects from environmental effects and their interactions.

7. AI Applications for Livestock Genetic Improvement

7.1 Milk Production

Machine learning methods have shown potential for predicting economically important dairy traits, such as milk yield and fertility, using genetic and performance data. Artificial neural networks have displayed good ability to predict these traits, but their performance can differ depending on the trait and the dataset used (Beskorovajni *et al.*, 2021). Studies comparing machine learning with traditional genomic prediction methods also show that no single method performs best in all situations. For example, Gradient Boosting gave the most accurate predictions of sire conception rate among the methods tested, performing better than GBLUP, Bayes B, Random Forest, and deep learning methods (Abdollahi-Arpanahi *et al.*, 2020). Combining automated activity monitoring with health, environmental, and farm information may provide additional information for predicting economically important traits.

7.2 Growth Traits

Machine-learning approaches have been applied to the genomic prediction of growth traits, including body weight and carcass characteristics. Genomic prediction of these traits involves high-dimensional SNP datasets, in which machine-learning methods can assist in identifying informative genomic markers and modelling complex genetic relationships. In Brahman cattle, Li *et al.* (2018) evaluated Random Forest, Gradient Boosting Machine, and XGBoost for identifying informative SNP subsets for genomic prediction of body weight. Random Forest and particularly Gradient Boosting Machine were effective in identifying informative SNP subsets, and prediction using selected SNP subsets performed better than prediction using evenly spaced SNP subsets. In Chinese Simmental beef cattle, Liang *et al.* (2021) evaluated support vector regression, kernel ridge regression, random forest, and AdaBoost.RT against GBLUP for genomic prediction of live weight, carcass weight, and eye muscle area. The machine-learning methods outperformed GBLUP in their study, although the magnitude of improvement varied among methods and traits. These findings indicate that machine-learning approaches can complement conventional genomic prediction by identifying informative genomic patterns and reducing the dimensionality of high-density SNP data; however, results from individual populations should not be interpreted as evidence of a universal advantage over GBLUP.

Thus, AI-based approaches may contribute to the genetic improvement of growth traits through genomic feature selection, dimensionality reduction, and prediction of genetic merit, but their relative advantage over conventional methods such as GBLUP remains dependent on the trait, population, genetic architecture, training population, and validation strategy.

7.3 Disease Resistance

Disease resistance is an important breeding objective in livestock because health disorders can reduce productivity, animal welfare, and longevity. Genomic selection can be used to estimate

genetic susceptibility to diseases and health disorders such as mastitis, ketosis, and lameness. However, the generally low heritability of many health traits and their strong environmental influences can make accurate genetic and genomic prediction challenging. AI and machine-learning methods offer opportunities to integrate high-dimensional genomic data with health phenotypes generated from sensors, imaging, and metabolic measurements to identify complex associations related to disease susceptibility and resistance. For example, machine-learning approaches have been used to classify mastitis susceptibility using high-dimensional genomic data and to identify SNPs associated with the predicted phenotype (Kotlarz *et al.*, 2024), while automated phenotyping can provide repeated and objective indicators of health status. Integrating these complementary data sources may provide more comprehensive information on disease-related phenotypes and potentially improve their utility for genetic evaluation. However, disease occurrence is strongly influenced by management, environmental conditions, and pathogen exposure. Therefore, accurate prediction of disease status does not necessarily imply accurate prediction of genetic resistance, because disease expression is influenced by both genetic and environmental factors.

7.4 Fertility

Fertility is an economically important but challenging breeding objective in livestock, particularly in dairy cattle, because reproductive traits generally have low heritability and are influenced by sex, age, management, nutrition, and environmental conditions. Nevertheless, genetic improvement is possible through direct and genomic selection, particularly when genetic information is combined with accurate reproductive records and technologies such as artificial insemination. Genomic prediction and machine-learning approaches have therefore been investigated to improve the prediction of reproductive performance. For example, Alves *et al.* (2021) evaluated support vector regression, Bayesian regularized artificial neural networks, and random forest alongside GBLUP and Bayesian models for reproductive traits in Nellore cattle, including age at first calving, scrotal circumference, early pregnancy, and stayability. Support vector regression showed slightly higher predictive accuracy than the parametric genomic prediction methods for the traits evaluated, although the relative performance varied among traits. Beyond genomic prediction, transcriptomics, proteomics, and metabolomics can provide additional biological information for identifying molecular markers or biomarkers associated with reproductive performance.

In dairy cattle, transcriptomic data from endometrial samples have been combined with machine-learning methods to identify gene-expression biomarkers predictive of pregnancy following artificial insemination (Hoorn *et al.*, 2024), demonstrating the potential of integrating molecular profiles with reproductive phenotypes. Such approaches may provide additional indicators for earlier identification of animals with favourable reproductive potential. However, before these

indicators can be incorporated into genetic selection programmes, their repeatability, genetic association with the target fertility trait, and predictive performance across independent populations and environments need to be established. Thus, AI and multi-omics approaches may complement conventional genetic evaluation by providing additional biological information, but their value for genetic improvement ultimately depends on demonstrated genetic relevance, robust validation, and their ability to contribute to reliable selection decisions.

8. Multi-Omics, Selection and Future Perspectives

The integration of genomics, phenomics and environmental information provides a multidimensional basis for precision livestock breeding. Genomics captures inherited variation, phenomics describes expressed traits, and enviromics defines the environments in which genetic potential is expressed. AI can facilitate the integration of heterogeneous data sources, particularly for complex traits influenced by multiple genetic and environmental factors. Multi-omics approaches can further link genetic variation with gene regulation, proteins, metabolites, and phenotypic expression, thereby improving the biological understanding and potentially enhancing the prediction of complex traits. However, the high dimensionality of multi-omics data relative to the typically limited sample sizes in livestock populations increases the risk of overfitting, making appropriate feature selection, regularization, and independent validation essential.

AI can also support selection and mating decisions by integrating genomic, phenotypic, and other relevant information to rank animals and evaluate potential mating strategies. Such approaches can potentially help optimise expected genetic gain while controlling inbreeding and maintaining genetic diversity. AI-derived predictions may help reduce the generation interval when reliable early-life information can be used to estimate the later genetic merit of animals, enabling selection at a younger age. Thus, the value of AI should not be judged solely by predictive accuracy but also by its contribution to selection efficiency and realized genetic response.

Future livestock breeding is likely to increasingly integrate AI with established genetic evaluation methods, large-scale digital phenotyping, and genetic and environmental information, alongside emerging approaches such as transfer learning and foundation models. Because livestock breeding populations are often relatively small and may have distinct genetic structures, the generalizability of these methods needs to be carefully evaluated using data from diverse populations, generations, and environments (Morota *et al.*, 2026).

9. Limitations and Conclusions

The use of AI for genetic improvement in livestock is affected by data quality, limited population size, overfitting, limited generalisability across populations, and computational requirements. Missing records, measurement errors, differences in trait-recording procedures, genotyping

errors, and differences between animal populations can reduce the reliability of predictions. In addition, randomly dividing closely related animals from the same dataset into training and testing groups may result in an unrealistically high estimate of model performance.

Complex AI models may achieve high predictive accuracy without providing a clear biological interpretation. Therefore, associations identified by AI should not automatically be treated as causal relationships. AI models should be compared with established quantitative genetic methods such as GBLUP and validated using independent populations, generations, and environments. The economic value of AI must also be considered. High computational costs, the cost of data collection, and the need for technical expertise may make these methods difficult to implement in small breeding programmes. AI also does not eliminate conventional breeding concerns such as inbreeding and the loss of genetic diversity. Thus, artificial intelligence is best regarded as a computational component of modern livestock genetic improvement. Its most immediate contribution is likely to be the generation of objective and scalable phenotypes from images, video, and sensor-based technologies.

In the longer term, AI may facilitate the integration of genetic, animal performance, environmental, and multi-omics data to support more comprehensive genetic evaluation and selection decisions. The future of livestock breeding is therefore likely to involve AI integrated with quantitative genetics, genomics, phenomics, and environmental data, with the ultimate goal of improving selection accuracy and achieving sustainable genetic progress in production, health, fertility, efficiency, longevity, and environmental resilience.

References

1. Abdollahi-Arpanahi, R., Gianola, D., & Peñagaricano, F. (2020). Deep learning versus parametric and ensemble methods for genomic prediction of complex phenotypes. *Genetics Selection Evolution*, 52, 12. <https://doi.org/10.1186/s12711-020-00531-z>
2. Alves, A. A. C., Espigolan, R., Bresolin, T., Costa, R. M., Fernandes Júnior, G. A., Ventura, R. V., Carneiro, R., & Albuquerque, L. G. (2021). Genome-enabled prediction of reproductive traits in Nelore cattle using parametric models and machine learning methods. *Animal Genetics*, 52(1), 32–46. <https://doi.org/10.1111/age.13021>
3. Beskorovajni, R., Jovanović, R., Pezo, L., Popović, N., Tolimir, N., Mihajlović, L., & Surlan-Momirović, G. (2021). Mathematical modeling for genomic selection in Serbian dairy cattle. *Genetika-Belgrade*, 53(3), 1105–1115.
4. Bewley, J. M., Peacock, A. M., Lewis, O., Boyce, R. E., Roberts, D. J., Coffey, M. P., Kenyon, S. J., & Schutz, M. M. (2008). Potential for estimation of body condition scores in dairy cattle from digital images. *Journal of Dairy Science*, 91(9), 3439–3453. <https://doi.org/10.3168/jds.2007-0836>

5. Boichard, D., Ducrocq, V., Croiseau, P., & Fritz, S. (2016). Genomic selection in domestic animals: Principles, applications and perspectives. *Comptes Rendus Biologies*, 339(7–8), 274–277. <https://doi.org/10.1016/j.crvi.2016.04.007>
6. Chafai, N., Hayah, I., Houaga, I., & Badaoui, B. (2023). A review of machine learning models applied to genomic prediction in animal breeding. *Frontiers in Genetics*, 14, 1150596.
7. Garrick, D. J., & Golden, B. L. (2009). Producing and using genetic evaluations in the United States beef industry of today. *Journal of Animal Science*, 87(suppl_14), E11–E18. <https://doi.org/10.2527/jas.2008-1431>
8. Hay, E. H. (2024). Machine learning for the genomic prediction of growth traits in a composite beef cattle population. *Animals*, 14(20), 3014.
9. Hoorn, Q. A., Rabaglino, M. B., Amaral, T. F., Maia, T. S., Yu, F., Cole, J. B., & Hansen, P. J. (2024). Machine learning to identify endometrial biomarkers predictive of pregnancy success following artificial insemination in dairy cows. *Biology of Reproduction*, 111(1), 54–62. <https://doi.org/10.1093/biolre/ioae052>
10. Iqbal, F. (2026). Role of artificial intelligence and smart technologies in improving production of ruminant livestock. *Applied Animal Science Bulletin*, 3(1), 71–79. <https://doi.org/10.22194/AASB/26.1058>
11. Kotlarz, K., Mielczarek, M., Biecek, P., Wojdak-Maksymiec, K., Suchocki, T., Topolski, P., Jagusiak, W., & Szyda, J. (2024). An explainable deep learning classifier of bovine mastitis based on whole-genome sequence data—Circumventing the $p \gg n$ problem. *International Journal of Molecular Sciences*, 25(9), 4715. <https://doi.org/10.3390/ijms25094715>
12. Lee, H.-J., Lee, J. H., Gondro, C., Koh, Y. J., & Lee, S. H. (2023). DeepGBLUP: Joint deep learning networks and GBLUP framework for accurate genomic prediction of complex traits in Korean native cattle. *Genetics Selection Evolution*, 55, 56.
13. Li, B., Zhang, N., Wang, Y.-G., George, A. W., Reverter, A., & Li, Y. (2018). Genomic prediction of breeding values using a subset of SNPs identified by three machine learning methods. *Frontiers in Genetics*, 9, 237. <https://doi.org/10.3389/fgene.2018.00237>
14. Liang, M., Miao, J., Wang, X., Chang, T., An, B., Duan, X., Xu, L., Gao, X., Zhang, L., Li, J., & Gao, H. (2021). Application of ensemble learning to genomic selection in Chinese Simmental beef cattle. *Journal of Animal Breeding and Genetics*, 138(3), 291–299. <https://doi.org/10.1111/jbg.12514>
15. Meuwissen, T. H. E., Hayes, B. J., & Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157, 1819–1829.

16. Montes, M. E., Menezes, G. L., Reinemann, D. J., Hernandez, L. L., & Dórea, J. R. R. (2026). Deep learning and image processing for high-throughput udder phenotyping in dairy cows. *Journal of Dairy Science*, 109(7), 7114–7131. <https://doi.org/10.3168/jds.2025-28008>
17. Morota, G., Lee, H.-J., Lee, S.-H., Park, H.-B., & Onogi, A. (2026). Artificial intelligence in animal breeding and genetics: Applications, opportunities, and challenges. *Animal Frontiers*, 16, 99–107. <https://doi.org/10.1093/af/vfaf058>
18. Nayak, S., Jena, L., Palai, P., Mishra, S., & Swain, M. K. (2024). Application of machine learning in the analysis and prediction of animal disease. In *Sustainable farming through machine learning* (pp. 207–219). CRC Press.
19. Nayeri, S., Sargolzaei, M., & Tulpan, D. (2019). A review of traditional and machine learning methods applied to animal breeding. *Animal Health Research Reviews*, 20(1), 31–46. <https://doi.org/10.1017/S1466252319000148>
20. Neethirajan, S., & Kemp, B. (2021). Digital phenotyping in livestock farming. *Animals*, 11(7), 2009.
21. Panigrahi, M., Rajawat, D., Nayak, S. S., Bose, A., Bharia, N., Singh, S., ... & Dutt, T. (2025). Advancements in animal breeding: From Mendelian genetics to machine learning. *International Journal of Molecular Sciences*, 26(23), 11352. <https://doi.org/10.3390/ijms262311352>
22. Pedrosa, V. B., Chen, S. Y., Gloria, L. S., Doucette, J. S., Boerman, J. P., Rosa, G. J., & Brito, L. F. (2024). Machine learning methods for genomic prediction of cow behavioral traits measured by automatic milking systems in North American Holstein cattle. *Journal of Dairy Science*, 107(7), 4758–4771.
23. Rohart, F., Gautier, B., Singh, A., & Lê Cao, K.-A. (2023). Dealing with dimensionality: The application of machine learning to multi-omics data. *Bioinformatics*, 39(2), btad021. <https://doi.org/10.1093/bioinformatics/btad021>
24. Silva, L. R. G., Barbosa, B. L., Silva, R. R. V. E., & Sarmiento, J. L. R. (2026). Body weight prediction and genetic evaluation in sheep: A review. *Ciência Rural*, 56(08), e20250267.
25. Spoliansky, R., Edan, Y., Parmet, Y., & Halachmi, I. (2016). Development of automatic body condition scoring using a low-cost 3-dimensional Kinect camera. *Journal of Dairy Science*, 99(9), 7714–7725. <https://doi.org/10.3168/jds.2015-10607>
26. Srivastava, S., Lopez, B. I., Kumar, H., Jang, M., Chai, H. H., Park, W., ... & Lim, D. (2021). Prediction of Hanwoo cattle phenotypes from genotypes using machine learning methods. *Animals*, 11(7), 2066. <https://doi.org/10.3390/ani11072066>

27. Tarjan, L., Šenk, I., Pracner, D., Rajković, D., & Štrbac, L. (2021, March). Possibilities for applying machine learning in dairy cattle breeding. In *2021 20th International Symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1–6). IEEE.
28. VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*, 4414–4423. <https://doi.org/10.3168/jds.2007-0980>
29. Ventura, R. V., e Silva, F. F., Yáñez, J. M., & Brito, L. F. (2020). Opportunities and challenges of phenomics applied to livestock and aquaculture breeding in South America. *Animal Frontiers: The Review Magazine of Animal Agriculture*, *10*(2), 45.
30. Wadood, A. A., Bordbar, F., & Zhang, X. (2025). Integrating omics approaches in livestock biotechnology: Innovations in production and reproductive efficiency. *Frontiers in Animal Science*, *6*, 1551244. <https://doi.org/10.3389/fanim.2025.1551244>
31. Wang, X., Shi, S., Wang, G., Luo, W., Wei, X., Qiu, A., ... & Ding, X. (2022). Using machine learning to improve the accuracy of genomic prediction of reproduction traits in pigs. *Journal of Animal Science and Biotechnology*, *13*(1), 60. <https://doi.org/10.1186/s40104-022-00708-0>
32. Weigel, K. A. (2017). Genomic selection of dairy cattle: A review of methods, strategies, and impact. *Journal of Animal Breeding and Genomics*, *1*(1), 1–15.
33. Zhang, L., Zhang, M., Wang, X., Chen, C., Wang, S., Zhao, Y., ... & Liu, Y. (2026). From linear models to deep learning: Statistical advances in genomic selection for animal breeding. *Journal of Animal Science and Biotechnology*, *17*(1), 159. <https://doi.org/10.1186/s40104-026-01477-w>
34. Zhao, W., Lai, X., Liu, D., Zhang, Z., Ma, P., Wang, Q., ... & Pan, Y. (2020). Applications of support vector machine in genomic prediction in pig and maize populations. *Frontiers in Genetics*, *11*, 598318.

AI-BASED HEALTH MONITORING OF FARM ANIMALS

Sunil Kumar*, Yogesh C. Bangar, Anil Chitra and Nitesh Jat

Department of Animal Genetics and Breeding,

Lala Lajpat Rai University of Veterinary and Animal Sciences (LUVAS), Hisar 125004 India

*Corresponding author E-mail: drsunil@luvas.edu.in

Abstract

Health monitoring is important for maintaining productivity, health and welfare of farm animals. Conventional health monitoring mainly depends on visual observation, clinical examination and production records. However, small changes in animal behaviour or physiological responses during the early stages of disease may not always be easily detected. Recent developments in artificial intelligence (AI) and precision livestock farming provide new opportunities for continuous monitoring of animals. Wearable sensors, cameras, microphones, automated farm systems and environmental sensors can generate information on animal activity, feeding, rumination, body temperature, posture, locomotion and environmental conditions. Machine-learning (ML) and deep-learning (DL) methods analyse these data and help in the early identification of changes associated with different health problems. This chapter provides an overview of the use of AI for health monitoring in farm animals, with emphasis on different sources of data and their applications in the monitoring of mastitis, lameness, respiratory diseases, metabolic disorders, reproductive and periparturient problems, heat stress and other welfare-related conditions. The chapter also explains the use of computer vision, acoustic monitoring and integration of behavioural, physiological, production and environmental data for animal health assessment. Finally, the chapter discusses future developments and the role of AI-based monitoring in early disease detection and better health management.

Keywords: Artificial Intelligence, Animal Health, Machine Learning, Precision Livestock Farming, Disease Detection.

1. Introduction

Animal health has an important influence on production, reproduction, longevity and welfare of farm animals. Diseases cause economic losses through mortality, treatment and culling as well as through reduction in milk yield, growth, fertility and feed efficiency. Infectious diseases may also spread rapidly within a herd or flock, and some diseases have implications for food safety and public health. Therefore, early identification of health problems is an important part of livestock management.

Health monitoring on livestock farms traditionally depends on observation by farm personnel, clinical examination, laboratory diagnosis and evaluation of production records. These methods are essential for diagnosis and treatment of disease. However, early stages of disease are sometimes associated with small changes in feed intake, activity, rumination, posture or body

temperature, which may not be easily noticed during routine observation. Individual monitoring also becomes difficult as herd or flock size increases.

Recent developments in artificial intelligence have increased the possibilities for early detection and risk assessment of diseases in livestock. In cattle, AI-based approaches are progressing from traditional sensor-based monitoring and machine-learning methods towards computer vision, deep learning and integration of behavioural, physiological and environmental data (Fang *et al.* 2026). The development of precision livestock farming has provided new methods for monitoring animals using sensors and other automated devices (Rutten *et al.*, 2013; Neethirajan, 2017). These technologies can continuously record behavioural, physiological, production and environmental information. AI and machine-learning methods can be used to analyse such data and identify changes that may be associated with health problems. This can help in selecting animals that require closer observation or clinical examination. Thus, AI-based monitoring can supplement routine farm observation and veterinary examination, particularly for the early detection of changes in animal health.

2. Artificial Intelligence in Animal Health Monitoring

Artificial intelligence (AI) includes computational methods that analyse data, recognize patterns and support prediction or decision-making. Machine learning (ML) is a major component of AI in which models learn relationships from existing data and use them to classify or predict new observations. Common ML methods used in animal-health applications include logistic regression, decision trees, random forests, support vector machines and gradient-boosting methods. Deep learning (DL), a subset of ML, is particularly useful for complex data such as images, videos, sounds and continuous sensor records. For example, convolutional neural networks (CNNs) can analyse images and videos, while other DL models can evaluate sequential or time-series data. An AI-based health-monitoring system generally involves data collection, preprocessing, model development, validation and generation of health alerts. Its usefulness depends not only on model accuracy but also on reliable validation and whether the alerts are timely, understandable and useful for farmers and veterinarians. AI-generated alerts should therefore support the already existing clinical examination and veterinary decision-making.

3. Sources of Data For AI-Based Health Monitoring

AI-based health monitoring depends on the continuous collection of reliable and biologically relevant data from animals and their environment. These data can be obtained from multiple sources (Rutten *et al.* 2013; Neethirajan, 2017) as described below:

3.1 Wearable Sensors

Wearable devices such as accelerometers, pedometers, smart collars, ear tags and leg-mounted sensors record activity, posture, locomotion, feeding, rumination, lying and standing. Changes in these behaviours frequently precede visible signs of disease. Reduced rumination and feed intake may accompany fever, mastitis, digestive disturbances or metabolic disease, while alter walking

and lie behaviour may indicate lameness or discomfort. AI algorithms can analyse continuous time-series data and identify deviations from the animal's individual baseline.

3.2 Physiological Sensors

Physiological monitoring includes monitoring of physiological parameters like body temperature, heart rate etc. Continuous temperature measurements may identify inflammatory or infectious processes earlier than occasional manual examination. The usefulness of physiological signals increases when they are interpreted along with behaviour and production. For example, elevated temperature combined with reduced activity and feed intake provides a more meaningful warning than a single abnormal parameter.

3.3 Computer Vision and Imaging

Computer vision provides contactless monitoring through two-dimensional cameras, depth cameras and 3-D imaging. Algorithms can identify animals and quantify posture, gait, body dimensions, body condition, feeding behaviour and social interactions. Video-based monitoring is particularly attractive in large groups because one system can observe several animals simultaneously. Thermal imaging can additionally detect changes in surface temperature associated with inflammation, fever or heat stress, although environmental conditions and anatomical location must be considered.

3.4 Acoustic and Environmental Sensors

Microphones can record coughing, vocalization and other sounds associated with respiratory disease, pain, stress or abnormal behaviour. AI-based acoustic analysis can separate biologically meaningful sounds from background farm noise and can also provide group-level surveillance in pigs, poultry and cattle. Environmental sensors measuring temperature, humidity, ventilation, gases and particulate matter provide contextual information because poor housing conditions can predispose animals to disease. Integrating animal and environmental measurements can therefore improve interpretation of health alerts.

4. Applications in Major Health Problems

4.1 Early Disease Detection

Early detection of disease is one of the major applications of AI in livestock health monitoring. Many diseases are associated with small changes in behaviour, physiological responses, production or environmental interaction. These changes may include reduced activity, altered feeding or rumination, changes in body temperature or unusual movement patterns. Recent developments have expanded from conventional sensor-based monitoring and ML methods to computer vision, DL and multimodal data fusion (Fang *et al.*, 2026). Using several types of information together may improve the detection of complex or subclinical health problems. Such systems can therefore serve as an early-warning tool and support timely clinical examination, diagnosis and appropriate management.

4.2 Mastitis Detection

Mastitis is among the most economically important diseases of dairy animals. Automated milking systems generate data on milk yield, electrical conductivity, milking duration and, in some systems, milk composition or somatic-cell-related indicators (Mitsunaga *et al.*, 2024). ML models can combine these variables with activity, rumination and previous health records to estimate mastitis risk. Such systems are most appropriately used for screening of animals. Animals identified as high risk should undergo clinical evaluation after identification and, further before making treatment decision, laboratory evaluation can also be done.

4.3 Lameness Monitoring

Lameness is a common health and welfare problem in dairy cattle and can affect milk production and reproductive performance. Visual scoring is commonly used to identify lame animals, but it requires time and the results may vary between observers. Automated lameness detection has been investigated using accelerometers, weighing platforms, acoustic sensors and computer vision. However, reported performance depends strongly on the reference standard used to classify lameness. These systems can identify changes in gait and may detect lameness at an early stage. Early detection allows farmers to examine affected animals and take timely measures to prevent the condition from becoming more severe (Siachos *et al.*, 2024).

4.4 Respiratory Diseases

Respiratory diseases in calves, cattle, pigs, and poultry cause considerable production losses. Coughing and altered vocalizations are key early indicators. AI systems continuously analyze acoustic data to detect respiratory issues early. Complementary metrics—body temperature, activity, feed intake, and ambient conditions—further refine health monitoring. Across livestock species, acoustic monitoring proves highly effective for early disease detection (Lagua *et al.*, 2023; Prosser *et al.*, 2025; Sharifuzzaman *et al.*, 2026). Combining these signals offers a robust early warning system, helping producers intervene promptly.

4.5 Metabolic and Digestive Disorders

Metabolic and digestive disorders are common in dairy cattle, particularly during periods of high production and around calving. Changes in feed intake, rumination, activity, milk composition and other behaviours may occur before occurrence of clear signs of disease. AI-based systems can analyse information from sensors, milk records, cameras and feeding behaviour to identify animals showing unusual changes. These tools may help in the early detection of conditions such as ketosis, acidosis and other digestive problems (Fang *et al.* 2026). Early alerts can help farmers identify animals that need attention and take timely management or treatment measures.

4.6 Reproductive and Periparturient Health

AI can be useful for monitoring different reproductive events in farm animals, such as estrus, pregnancy and calving (Giordano and Laplacette, 2026). Sensors can record changes in activity, rumination, body temperature and feeding behaviour, which may help in predicting these events.

Such information can also be used to identify animals that need more attention during pregnancy and around calving. For example, unusual changes in normal behaviour may be associated with dystocia, retained fetal membranes, metritis or other postpartum problems. However, many tools remain at an early stage and that biological and management factors influence model performance. Continuous monitoring can provide an early warning to farmers, especially when these changes occur outside routine observation periods.

4.7 Heat Stress and Welfare-related Problems

Heat stress is an important concern in livestock, particularly in hot and arid regions. It can reduce feed intake, milk production, growth and reproductive performance and may also affect animal health and welfare. Environmental sensors can continuously record temperature and humidity, while animal-based sensors can measure changes in body temperature, respiration, activity and rumination. AI models can analyse these measurements together and help identify animals showing signs of heat stress. These systems can support timely interventions such as provision of shade, cooling, altered feeding schedules and other management measures, thereby helping to reduce heat-related production losses and improve animal welfare (Rebez *et al.*, 2024; Emsen *et al.*, 2026).

5. Multimodal Data Integration and Predictive Analytics

A single sensor may not provide enough information to assess animal health correctly. For example, reduced activity for example, may result from disease, estrus, environmental conditions or normal individual variation. Therefore, information from different sources, such as activity sensors, milk records, body temperature, cameras, environmental conditions and health history, can be combined. When changes are observed in several indicators at the same time, disease alerts may become more reliable and false alerts can be reduced.

Predictive methods can also be used to identify animals that are at greater risk of developing disease in the future. Information such as age, parity, previous disease, production level, genetic background and current sensor records can be considered together for risk assessment. This can help farmers and veterinarians take preventive measures before the disease becomes severe. However, these prediction models need proper validation under actual farm conditions before their routine use.

6. Benefits of AI-based Health Monitoring

Continuous monitoring is an important advantage of automated health systems. These systems can work throughout the day and night and monitor many animals at the same time. Early detection of abnormal animals allows timely examination and treatment, which may improve recovery and help control the spread of infectious diseases. It can also support the careful use of antimicrobials by identifying animals that actually require further diagnosis and treatment.

Automated monitoring can also reduce the workload of farm staff by drawing attention to animals that need closer observation. Objective measurements reduce differences between

individual observers and provide continuous health records that can support veterinary decisions. At the herd level, these records can help identify changes in disease occurrence, environmental risks and overall management. Such technologies can also improve animal welfare by allowing earlier recognition of pain, discomfort and illness. Combining health information with production and reproduction records provides a better understanding of the overall effects of disease on animal performance and farm profitability.

7. Challenges and Limitations

Despite its usefulness, the application of AI for livestock health monitoring has several practical limitations. One of the major concerns is the quality of data used for developing prediction models. Model performance might reduce due to several factors such as missing observations, malfunctioning sensors, errors in animal identification and inaccurate disease records. Moreover, a model developed using data from a particular farm, breed or production environment may not perform equally well under different conditions. Therefore, testing these models on independent farms and animal populations is important before their wider application. The occurrence of false alerts is another concern. Frequent false-positive alerts increase unnecessary work for farmers and may gradually reduce their confidence in the monitoring system. On the other hand, false-negative results are more serious because affected animals may remain undetected and treatment may be delayed. Hence, evaluation of a health-monitoring system should not depend on overall accuracy alone. Measures such as sensitivity, specificity and predictive values, along with the practical consequences of incorrect alerts, should also be considered.

Cost and availability of infrastructure may also restrict the adoption of these technologies, particularly on small and medium-sized farms. Sensors and other monitoring devices require installation, regular maintenance, reliable power supply, connectivity and technical support. Another practical difficulty arises when equipment from different manufacturers uses different data formats, making it difficult to combine information within a single monitoring system. Issues related to ownership, privacy, security and sharing of farm data also need appropriate attention.

The usefulness of an AI system ultimately depends on whether its output can be understood and used by farmers and veterinarians. An alert is more useful when the user has some indication of why an animal has been identified as abnormal. Explainable AI approaches may help in this regard by indicating which measurements contributed most to a prediction. However, changes detected by sensors are often not specific to a particular disease. AI-generated alerts should therefore be interpreted together with clinical findings, disease history and farm-management conditions before decisions regarding treatment or other interventions are made.

8. Future Perspectives

AI-based health monitoring is likely to become more useful as sensors become affordable and easier to use under farm conditions. At present, continuous monitoring may be difficult on farms

with limited internet connectivity, power supply or technical facilities. Processing some of the sensor data directly at the farm may help overcome these limitations. Improvements in low-cost and energy-efficient sensors may also increase their use in different livestock production systems.

Future systems are expected to use information from several sources rather than depending on a single sensor. Activity, feeding behaviour, body temperature, images, sounds, environmental conditions and production records provide different information about the health of an animal. Their combined analysis may improve the identification of health problems, particularly when changes in one parameter alone are not sufficient to indicate disease. These records may also provide useful information on health and disease resistance for genetic evaluation.

Another useful approach could be to compare an animal with its own normal behaviour. Animals differ in their activity, feeding, rumination and physiological responses according to age, breed, production stage and health history. Therefore, the same threshold may not be suitable for every animal. Monitoring changes from the normal pattern of an individual animal may provide more meaningful health alerts.

The performance of AI models may also improve when information from different farms and production conditions is included during their development. Approaches such as federated learning may allow models to learn from data generated on different farms without transferring all farm data to a common location. Explainable AI may also be useful in showing farmers and veterinarians the main factors responsible for a particular health alert. However, the future usefulness of these technologies will depend mainly on their performance under actual farm conditions. Models with high accuracy during development may not always perform equally well on different breeds, farms or management systems. Therefore, greater attention should be given to field validation, cost, ease of use and the practical value of the information provided. Cooperation among animal scientists, veterinarians, farmers, engineers and data scientists will be important for developing AI tools that can be routinely used for livestock health management.

Conclusions

Artificial intelligence is becoming a useful tool for monitoring the health of farm animals. Technologies such as wearable sensors, cameras, microphones, thermal imaging and environmental sensors can continuously collect information on animal behaviour, body condition, activity and surroundings. AI can analyse these data and help identify animals that may be at risk of health problems such as mastitis, lameness, respiratory diseases, metabolic disorders and heat stress. An important advantage of AI is its ability to detect small changes in an animal before clear signs of disease appear. This can help farmers and veterinarians examine affected animals earlier and take timely preventive or treatment measures. However, the usefulness of these systems depends on good-quality data, reliable performance, reasonable cost and proper validation under farm conditions. AI should therefore be used as a supporting tool

along with farmers' experience and veterinary judgement. When used properly, these technologies can help improve animal health and welfare, reduce production losses and make livestock management more efficient.

References

1. Emsen, E., Odevci, B. B., Akgol, O., Korkmaz, M. K., & Yaganoglu, A. M. (2026). AI-enabled electromagnetic sensor architecture for precision welfare monitoring in small ruminant production systems. *Computers and Electronics in Agriculture*, 251, 112067.
2. Fang, J., Zhang, R., Li, X., Long, X., Jiang, Y., Wang, H., ... & Wang, H. (2026). Application of artificial intelligence in cattle disease diagnosis: A comprehensive review. *Artificial Intelligence in Agriculture*.
3. Giordano, J. O., & Laplacette, A. L. (2026). Present and future opportunities for artificial intelligence applications in dairy cattle reproduction. *Animal Frontiers*, 16(2), 86-98.
4. Lagua, E. B., Mun, H. S., Ampode, K. M. B., Kim, Y. H., & Yang, C. J. (2023). Artificial intelligence for automatic monitoring of respiratory health conditions in smart swine farming. *Animals*, 13(11), 1860.
5. Mitsunaga, T. M., Nery Garcia, B. L., Pereira, L. B. R., Costa, Y. C. B., da Silva, R. F., Delbem, A. C. B., & Dos Santos, M. V. (2024). Current trends in artificial intelligence and bovine mastitis research: A bibliometric review approach. *Animals*, 14(14), 2023.
6. Neethirajan, S. (2017). Recent advances in wearable sensors for animal health management. *Sensing and Bio-Sensing Research*, 12, 15-29.
7. Prosser, H. M., Bortoluzzi, E. M., Valeris-Chacin, R. J., Baker, E. C., & Scott, M. A. (2025). Application of artificial intelligence and machine learning in bovine respiratory disease prevention, diagnosis, and classification. *American Journal of Veterinary Research*, 86(S1), S22-S26.
8. Rebez, E. B., Sejian, V., Silpa, M. V., Kalaignazhal, G., Thirunavukkarasu, D., Devaraj, C., ... & Dunshea, F. R. (2024). Applications of artificial intelligence for heat stress management in ruminant livestock. *Sensors*, 24(18), 5890.
9. Rutten, C. J., Velthuis, A. G. J., Steeneveld, W., & Hogeveen, H. (2013). Sensors to support health management on dairy farms (invited review). *J Dairy Sci*, 96(4), 1928-1952.
10. Sharifuzzaman, M., Mun, H. S., Lagua, E. B., Hasan, M. K., Kang, J. G., Kim, Y. H., ... & Yang, C. J. (2026). Advances in Audio-Based Artificial Intelligence for Respiratory Health and Welfare Monitoring in Broiler Chickens. *AI*, 7(2), 58.
11. Siachos, N., Neary, J. M., Smith, R. F., & Oikonomou, G. (2024). Automated dairy cattle lameness detection utilizing the power of artificial intelligence; current status quo and future research opportunities. *The Veterinary Journal*, 304, 106091.

APPLICATION SECURITY CHALLENGES IN AI-ENABLED SOFTWARE SYSTEMS: RISKS, CONTROLS, AND SECURE DEVELOPMENT PRACTICES

Udaya Bhaskar Vemuri

Department of Security Analysis Application and DevSecOps Professional

Corresponding author E-mail: udv0507@gmail.com

Abstract

Artificial Intelligence (AI) is being added to business applications at a fast pace. Many systems now use AI to summarize documents, answer questions, generate code, review security findings, search internal knowledge, and automate routine work. These features can improve productivity, but they also change the application attack surface. An AI-enabled application may connect to APIs, databases, cloud services, source-code repositories, and business tools, often using service accounts or automated identities. This chapter explains common application-security risks in AI-enabled software systems using practical examples. The focus is on excessive permissions, sensitive-data exposure, prompt injection, insecure APIs, secrets, third-party integrations, software supply-chain risk, and AI agents that can take actions. It also shows how familiar Secure Software Development Lifecycle (Secure SDLC) and DevSecOps practices can be extended to AI applications without creating a completely separate security process. The central idea is simple: the AI model should not be treated as a trusted security boundary. Authentication, authorization, least privilege, data protection, secure APIs, human approval, testing, logging, and monitoring should be enforced by the surrounding application and platform. Organizations that apply these controls early can adopt AI more confidently while reducing the chance that a small design weakness becomes a larger security incident.

Keywords: Artificial Intelligence, Application Security, AI Agents, Secure Software Development, DevSecOps, API Security, Access Control, Data Protection

1. Introduction

Artificial Intelligence is no longer limited to research projects or experimental chatbots. Organizations are adding AI to customer portals, employee tools, development platforms, security systems, and everyday business workflows. A user may ask an AI assistant to summarize a policy, review a software vulnerability, prepare a report, search internal documents, or generate code. In more advanced use cases, an AI agent may also call an API, create a ticket, update a record, or trigger another automated process. [1, 2]

These capabilities are useful because they reduce manual work and make information easier to access. At the same time, they create security questions that are different from those in a traditional application. A normal application usually follows logic written directly by developers.

An AI-enabled system may interpret natural-language instructions, use generated output, retrieve external information, and decide which tool to call next. That flexibility can create unexpected paths through the application [1, 3]. For application-security teams, the important point is that the AI model is only one part of the system. The larger risk often comes from what is connected to it: service accounts, APIs, databases, secrets, cloud resources, third-party tools, and internal business applications. If those connections have broad permissions or weak authorization, an AI feature can increase the impact of a mistake or attack. [4, 5]

This chapter takes a practical AppSec view. It explains where AI changes the attack surface, which controls remain the same, and where additional safeguards are needed. The goal is not to slow AI adoption. The goal is to make sure that AI is introduced with the same discipline organizations already expect from important software systems. [4, 6]

2. How AI Changes the Application Attack Surface

Every new feature creates another area that needs to be protected. AI features are no different, but they often connect several parts of an environment that were previously separate. A simple chatbot may only send a prompt to a model. An enterprise assistant may also search documents, read customer records, call internal APIs, access source-code repositories, or perform actions in cloud platforms.

Table 1: AI-Enabled Application Components and Security Questions

Component	Security Question	Practical Control
User interface	Who is making the request?	Strong authentication and session controls
Application backend	Is the user allowed to perform the requested action?	Server-side authorization
AI model/service	What data and instructions are being sent?	Data minimization and prompt controls
Connected APIs/tools	What can the AI reach or change?	Least privilege and scoped permissions
Secrets/credentials	How are service identities protected?	Secrets manager and key rotation
Logs/monitoring	Can unusual activity be investigated?	Audit logging and alerting

Source: Author's own compilation based on the cited security frameworks.

Consider an internal assistant that begins as a tool for answering questions about company policies. The first version may only read a small set of approved documents. Later, the same assistant is connected to a ticketing system, an employee directory, and a customer database. The feature becomes more useful, but its security impact changes. The team now needs to understand which identity the assistant uses, what information it can retrieve, whether it can modify data, and what happens when a user tries to make it do something outside its intended purpose.

This is why AI threat modeling should focus on data flows and actions, not only on model behavior. The security team should identify where the request comes from, what the application sends to the model, what tools the model can call, which credentials those tools use, and where output is returned. These trust boundaries are often where practical security problems appear. [1, 4]

3. Common Application Security Risks in AI-Enabled Systems

3.1 Excessive Permissions

One of the most practical risks is giving an AI application more access than it needs. Broad permissions are sometimes used because they make integration easier. For example, an internal assistant may only need to read basic customer information, but the service account behind it may have permission to read, update, and delete customer records. If the application is compromised or the AI is manipulated, those extra permissions increase the possible impact. Least privilege should therefore be applied to every AI-connected identity. Read-only tasks should use read-only access, and separate functions should use separate service identities when possible. [4, 7]

3.2 Sensitive Data Exposure

AI tools encourage users to communicate in natural language, which makes it easy to share too much information. A developer may paste source code into an assistant. An employee may upload a customer spreadsheet. A support team may submit account details for summarization. The security concern is not only what the user intentionally shares. Retrieval systems may also bring sensitive information into a prompt when access controls are weak. Organizations should classify what data AI systems may process, use approved platforms, minimize unnecessary fields, mask sensitive values when possible, and prevent restricted information from being sent to services that are not authorized to handle it. [1, 8]

3.3 Prompt Injection and Untrusted Instructions

Prompt injection happens when crafted content attempts to influence the AI to ignore its intended instructions, reveal information, or perform an unsafe action. The instruction may come directly from a user, but it can also appear inside a webpage, document, email, or other content that the AI is asked to read. This is important because AI systems may have difficulty distinguishing data from instructions. The safest design is to assume that user and retrieved content are untrusted. Sensitive actions should be controlled by application logic, not by asking the model to decide whether an action is allowed. [3, 9]

3.4 Insecure APIs and Tool Connections

AI applications frequently depend on APIs. The model may call an internal service to retrieve data, create a ticket, send a notification, or start a workflow. These APIs still need the same protections used in any other application: authentication, authorization, input validation, rate

limiting, secure error handling, and logging. An internal API should not trust a request simply because it came from an AI component. The backend should verify the identity and permission of the user or service for each sensitive operation. [5, 10]

3.5 Secrets and Service Accounts

API keys, tokens, cloud credentials, and service-account passwords are often needed to connect AI applications to other systems. Hardcoding these values in source code or configuration files creates avoidable risk. Secrets should be stored in an approved secrets-management platform and retrieved at runtime. Access should be limited, credentials should be rotated, and unused secrets should be removed. Teams should also avoid using one powerful shared account for unrelated AI capabilities because it becomes difficult to control or investigate individual actions. [4, 6]

3.6 AI Agents Performing High-Impact Actions

The risk increases when AI moves from answering questions to taking actions. An AI agent may send email, change cloud resources, update a database, modify code, or approve a workflow. A wrong answer is inconvenient; a wrong action can directly affect production systems or business data. Organizations should therefore separate low-risk automation from high-impact actions. Drafting a response may be automatic, while deleting data, changing permissions, modifying production infrastructure, or releasing software should require an additional policy check or human approval. [1, 3]

Table 2: Common AI Application Security Risks and Practical Mitigations

Risk	Possible Impact	Practical Mitigation
Excessive permissions	Large blast radius if the AI is misused	Least privilege and scoped service identities
Sensitive data exposure	Confidential information leaves approved boundaries	Data classification, minimization, masking
Prompt injection	Manipulated model or tool behavior	Treat retrieved content as untrusted and enforce policy outside the model
Insecure APIs	Unauthorized access or actions	Strong authentication and server-side authorization
Exposed secrets	Compromise of connected systems	Secrets management, rotation, and scanning
Autonomous high-risk actions	Operational or business impact	Human approval and policy gates

Source: Author's own compilation based on OWASP, NIST, and secure-software guidance.

4. Software Supply Chain and Third-Party Risk

AI-enabled applications are built from many components. In addition to normal application libraries, they may use AI SDKs, orchestration frameworks, vector databases, plugins, model files, browser tools, and external APIs. Each dependency creates another place where vulnerabilities, malicious code, or insecure defaults may enter the environment. [4, 6]

Software Composition Analysis remains useful for identifying vulnerable packages and maintaining visibility into dependencies. A Software Bill of Materials can also help teams understand which components are included in an application. For AI-specific components, teams should record the source, version, owner, and intended use of important models, libraries, and plugins. Components downloaded from unknown sources should not automatically be trusted just because they are popular or easy to install. Third-party AI services need a different but related review. Before sensitive business data is sent to an external provider, organizations should understand how the service handles retention, access, geographic processing, incident response, and model training. Security, privacy, legal, and business owners may all need to participate when the use case involves confidential or regulated data. [1, 8]

5. Extending Secure SDLC and DevSecOps to AI Applications

Organizations do not need to throw away their existing Secure SDLC because an application includes AI. The better approach is to extend familiar controls so they also cover prompts, model connections, retrieval sources, AI tools, and automated actions. NIST's Secure Software Development Framework emphasizes integrating security throughout development rather than treating it as a final release activity. The same idea works well for AI-enabled software. [4]

5.1 Planning and Threat Modeling

Security should begin when the use case is defined. Teams should document what problem the AI is solving, what data it will process, which users will use it, what systems it can reach, and whether it can perform actions. Threat modeling should identify trust boundaries, sensitive assets, abuse cases, and the effect of a compromised service identity. A low-risk document summarizer does not need the same controls as an AI agent that can change production systems.

5.2 Secure Architecture and Access Control

Architecture reviews should show the full request path from user to model to connected tools. Authorization should be enforced by trusted backend services. If a user is not allowed to view a record through the normal application, the AI assistant should not create a new path around that restriction. Role-based or attribute-based access controls should continue to apply when data is retrieved for an AI response. [5, 7]

5.3 Secure Coding and Automated Testing

Traditional security testing is still necessary. AI applications can contain SQL injection, cross-site scripting, insecure deserialization, broken access control, exposed secrets, vulnerable

dependencies, and security misconfiguration. Static Application Security Testing, Software Composition Analysis, secret scanning, API testing, Dynamic Application Security Testing, and code review should continue to run where they make sense. These checks can be placed in CI/CD pipelines so developers receive feedback before release. [4, 6]

5.4 AI-Specific Security Testing

Teams should add tests that reflect how AI can be misused. This includes prompt injection, attempts to retrieve restricted data, unsafe tool calls, unexpected file types, manipulated retrieved content, attempts to bypass approval steps, and failure scenarios. Testing should include both normal user behavior and deliberately malicious instructions. The purpose is not to prove that the model will never make a mistake. It is to confirm that surrounding controls prevent a model mistake from turning into an unauthorized action. [3, 9]

5.5 Release and Change Control

AI applications can change even when the application code does not. A model version may change, a prompt may be updated, a new data source may be connected, or a tool may receive new permissions. These changes should be visible and reviewed. Teams should record important model and prompt versions, define owners, maintain rollback options, and use stronger approval for changes that increase access or automated authority. This keeps AI changes inside the same governance discipline used for other production software.

Table 3: Secure AI Application Development Activities

Stage	Primary Objective	Application Security Activities
Planning	Define purpose and boundaries	Data classification, threat modeling, abuse cases
Design	Build secure architecture	Identity, authorization, least privilege, secure APIs
Development	Implement safely	Secure coding, dependency review, secrets management
Testing	Find weaknesses before release	SAST, SCA, API testing, prompt-injection and tool-abuse tests
Deployment	Release with control	Configuration review, approvals, rollback planning
Operations	Detect misuse and failures	Logging, monitoring, incident response, vulnerability management
Improvement	Adapt to new threats	Permission review, updated tests, lessons learned

Source: Author's own compilation based on Secure SDLC, DevSecOps, and AI-security guidance.

6. Logging, Monitoring, and Incident Response

AI applications need enough logging to answer basic investigation questions: who made the request, which identity was used, what data source was accessed, which tool was called, what action was attempted, whether it succeeded, and whether approval was required. This is especially important for agents because actions may be spread across several systems.

Logging also needs restraint. Prompts and responses can contain confidential data, credentials, or personal information. Storing every prompt without filtering may create a new sensitive-data repository. Teams should decide what is necessary for audit and troubleshooting, mask sensitive values where practical, restrict log access, and define retention periods. [1, 8]

Monitoring should focus on behavior that is unusual for the application. Examples include a sudden increase in data retrieval, repeated attempts to access restricted records, an AI agent calling a tool it rarely uses, large numbers of failed authorization requests, or service accounts performing actions outside normal hours. These signals can support faster investigation and help identify whether the issue came from a user, the application, a compromised credential, or unexpected AI behavior. [6, 7]

7. Practical Example: Securing an Internal AI Support Assistant

Consider a company building an AI assistant for employees who support customers. The assistant can search approved knowledge articles and retrieve limited account details so employees can answer questions faster. The first design uses a shared service account with broad database access because it is easy to implement.

During the security review, the team identifies that the assistant only needs a small set of customer fields and does not need permission to update records. The service identity is changed to read-only access, and the backend is updated to check the employee's authorization before returning account information. Customer secrets and authentication data are excluded from the AI prompt. The team also tests prompt injection by entering instructions that ask the assistant to reveal restricted information or call unsupported tools. The model sometimes attempts to follow the instruction, but the backend authorization layer blocks the request. This is an important result: the security control does not depend on the model behaving perfectly.

Before release, the organization enables audit logging, adds monitoring for unusual data access, stores API credentials in a secrets manager, and requires human confirmation before any future feature is allowed to modify customer records. The assistant remains useful, but the possible impact of a mistake is significantly reduced. This example shows why AI security should be treated as application security around an unpredictable component. The model can assist with decisions, but identity, permission, data protection, and high-impact actions remain under the control of trusted software and accountable people. [3–5]

8. Best Practices for Organizations

- Start security reviews when the AI use case is being planned, before data access and integration decisions are difficult to change.
- Document every system, API, data source, and action that the AI application can reach.
- Use least privilege for users, service accounts, APIs, and agent tools, and review permissions when features change.
- Keep authorization in the application or backend service instead of relying on the model to decide what is allowed.
- Protect secrets through approved secrets-management systems and include secret scanning in development workflows.
- Test both traditional application vulnerabilities and AI-specific abuse cases such as prompt injection and unsafe tool calls.
- Require human approval or policy gates for actions that can materially affect production systems, customer data, money, permissions, or external communications.
- Log important AI actions, monitor unusual behavior, and make sure incident-response teams can trace activity across connected systems.

These practices are most effective when they are part of existing software-security and risk processes. A repeatable control that every AI project follows is usually more sustainable than relying on individual teams to make separate decisions each time. [1, 4, 6]

9. Future Directions

AI agents will continue to become more capable and more connected to business systems. As that happens, identity and authorization will become even more important. Organizations will need clearer ways to define what an AI agent is allowed to do, how its permissions are approved, and how actions can be traced back to the user or business process that initiated them.

Security testing will also need to become more continuous. New models, tools, plugins, and attack techniques can change the behavior of an AI application after the first release. Teams that build AI security into Secure SDLC, CI/CD, monitoring, and incident response will be better prepared than teams that treat AI review as a one-time checklist. [1, 3, 4]

Conclusion

AI is changing how applications are built and how users interact with software. The biggest security challenge is not only that AI can produce an incorrect response. It is that AI-enabled applications may be connected to sensitive data, powerful APIs, service accounts, and automated actions. Many of the controls needed to manage this risk are already familiar to application-security teams. Strong authentication, server-side authorization, least privilege, secure APIs, secrets management, dependency review, vulnerability testing, logging, and monitoring remain essential. AI-specific risks such as prompt injection and unsafe tool use should be added to those existing practices rather than treated as a completely separate security program.

The most important design principle is to avoid treating the model as a security boundary. Models can help interpret requests and generate useful output, but trusted application logic should control access, sensitive actions, and policy decisions. High-impact operations should include additional approval when appropriate. Organizations do not need to choose between AI adoption and application security. When security is included from planning through operations, AI systems can be useful without receiving unnecessary trust or access. The result is a more practical path to adopting AI while protecting users, data, applications, and business operations.

References

1. Tabassi, E. (2023). *Artificial intelligence risk management framework (AI RMF 1.0)* (NIST AI 100-1). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.100-1>
2. Autio, C., Schwartz, R., Dunietz, J., Jain, S., Stanley, M., Tabassi, E., Hall, P., & Roberts, K. (2024). *Artificial intelligence risk management framework: Generative artificial intelligence profile* (NIST AI 600-1). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.AI.600-1>
3. OWASP Foundation. (2024). *OWASP Top 10 for LLM applications 2025*. OWASP GenAI Security Project. <https://genai.owasp.org/resource/owasp-top-10-for-llm-applications-2025/>
4. Scarfone, K., Souppaya, M., & Dodson, D. (2022). *Secure software development framework (SSDF) version 1.1: Recommendations for mitigating the risk of software vulnerabilities* (NIST SP 800-218). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-218>
5. OWASP Foundation. (2023). *OWASP Top 10 API security risks—2023*. OWASP API Security Project. <https://owasp.org/API-Security/editions/2023/en/0x11-t10/>
6. Pascoe, C., Quinn, S., & Scarfone, K. (2024). *The NIST cybersecurity framework (CSF) 2.0* (NIST CSWP 29). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.CSWP.29>
7. Rose, S., Borchert, O., Mitchell, S., & Connelly, S. (2020). *Zero trust architecture* (NIST SP 800-207). National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.SP.800-207>
8. International Organization for Standardization. (2023). *ISO/IEC 42001:2023, Information technology—Artificial intelligence—Management system*. <https://www.iso.org/standard/42001.html>
9. MITRE. (n.d.). *MITRE ATLAS*. Retrieved September 8, 2026, from <https://atlas.mitre.org/>
10. OWASP Foundation. (2025). *OWASP application security verification standard (ASVS) 5.0.0*. <https://owasp.org/www-project-application-security-verification-standard/>

ARTIFICIAL INTELLIGENCE IN VETERINARY DIAGNOSTIC IMAGING: CURRENT APPLICATIONS, CHALLENGES AND FUTURE PERSPECTIVES

Antra Ken*¹, Sangeet Rana¹ and Aashish²

¹Department of Veterinary Surgery and Radiology,

²Department of Veterinary Gynaecology and Obstetrics,

M. R. Veterinary College & Research Center, Jhajjar, Haryana, India

*Corresponding author E-mail: invk.antraken28@gmail.com

Abstract

Artificial intelligence (AI) is becoming an important research and decision-support technology in veterinary diagnostic imaging. The increasing availability of digital radiography, ultrasonography, computed tomography (CT), magnetic resonance imaging (MRI), cone-beam CT and other digital imaging modalities has created large volumes of image data that can be analysed computationally. Machine learning (ML) and deep learning (DL), particularly convolutional neural networks, have been investigated for image classification, lesion detection, segmentation, measurement, image-quality assessment and workflow support. Recent veterinary literature demonstrates growing applications across radiography, CT, MRI and ultrasonography, with dogs representing a major proportion of published work. [1–5]

In veterinary radiology, AI may assist with thoracic lesion classification, cardiac size assessment, fracture detection, automated segmentation, radiomics and quantitative image analysis. For example, published studies have explored deep-learning systems for canine thoracic radiographs, automated vertebral heart size estimation and automated detection or segmentation of orthopaedic abnormalities. These developments are relevant to veterinary surgery because quantitative imaging and three-dimensional anatomical information may also support preoperative planning. [6–9]

However, the translation of AI from research into routine veterinary practice remains challenging. Important limitations include relatively small and heterogeneous datasets, inconsistent image acquisition, limited expert annotations, class imbalance, spectrum bias, insufficient external validation, model interpretability and uncertainty about performance across breeds, institutions and equipment. Professional guidance from the American College of Veterinary Radiology and European College of Veterinary Diagnostic Imaging emphasizes transparency, secure data handling, error reporting, post-implementation monitoring and continued involvement of qualified veterinary professionals. [4,10]

The future of veterinary imaging AI is likely to involve multimodal systems, automated reporting, quantitative radiomics, image-guided surgery, three-dimensional planning and

integration with clinical records. AI should be regarded as a clinical decision-support technology that complements rather than replaces veterinary expertise. Rigorous validation, appropriate governance and continuous monitoring will be essential for safe and useful implementation.

Keywords: Artificial Intelligence, Machine Learning, Deep Learning, Veterinary Radiology, Diagnostic Imaging, Radiography, Computed Tomography, Magnetic Resonance Imaging, Ultrasonography, Radiomics, Veterinary Surgery.

1. Introduction

Diagnostic imaging is a fundamental component of modern veterinary practice. Radiography is widely used for thoracic, abdominal and musculoskeletal assessment, while ultrasonography provides real-time soft-tissue evaluation. CT is particularly useful for complex anatomy, trauma, orthopaedic disease and oncological assessment, and MRI provides high soft-tissue contrast for neurological, musculoskeletal and other disorders. The expanding use of digital imaging has produced increasingly large collections of images that can be stored, shared and analysed.

Interpretation of veterinary images requires knowledge of anatomy, pathology, imaging physics, positioning, technical factors and clinical context. The process can be time-consuming, and observer variation may occur. These characteristics have encouraged investigation of computational approaches that can support image interpretation.

Artificial intelligence is an umbrella term for computational approaches designed to perform tasks associated with learning, pattern recognition, prediction or decision support. Machine learning is a subset of AI in which algorithms learn relationships from data, while deep learning uses multilayer neural networks to learn increasingly complex representations. Convolutional neural networks (CNNs) have become particularly important in image analysis. [1,2]

Veterinary imaging AI is still developing. A 2022 literature review identified fewer than 40 peer-reviewed publications using machine learning for imaging-associated veterinary tasks at that time and highlighted the need for larger datasets, high-quality ground-truth labels and closer collaboration between veterinary and computational experts. [1] More recent reviews indicate rapid expansion of the field across radiography, MRI, CT, cone-beam CT and ultrasound. [4,5]

2. Artificial Intelligence, Machine Learning and Deep Learning

2.1 Artificial Intelligence

AI includes a broad range of computational methods. In diagnostic imaging, AI may be used for classification, detection, segmentation, measurement, quality control, workflow prioritization and decision support. The output may be a probability, a class label, a highlighted region or a quantitative measurement.

2.2 Machine Learning

Machine learning systems learn patterns from examples. In supervised learning, images are paired with labels or reference measurements. The model is trained on one portion of the dataset

and evaluated on previously unseen images. Common tasks include binary classification, multiclass classification, object detection and regression.

2.3 Deep Learning

Deep learning uses neural networks with multiple computational layers. CNNs are especially suitable for image analysis because convolutional filters can learn spatial patterns. Depending on the task, architectures may be used for classification, detection or segmentation. Transfer learning can be useful when veterinary datasets are relatively small, because a model pretrained on a larger dataset can be adapted to a veterinary imaging task. [2,3]

2.4 Typical AI Imaging Workflow

A simplified workflow for development and clinical translation is:

- Image acquisition and secure data storage
- Data cleaning and quality control
- Annotation or creation of reference labels
- Pre-processing and normalization
- Training and internal validation
- Testing on unseen data
- External or multi-centre validation
- Clinical evaluation and workflow integration
- Post-implementation monitoring and error reporting

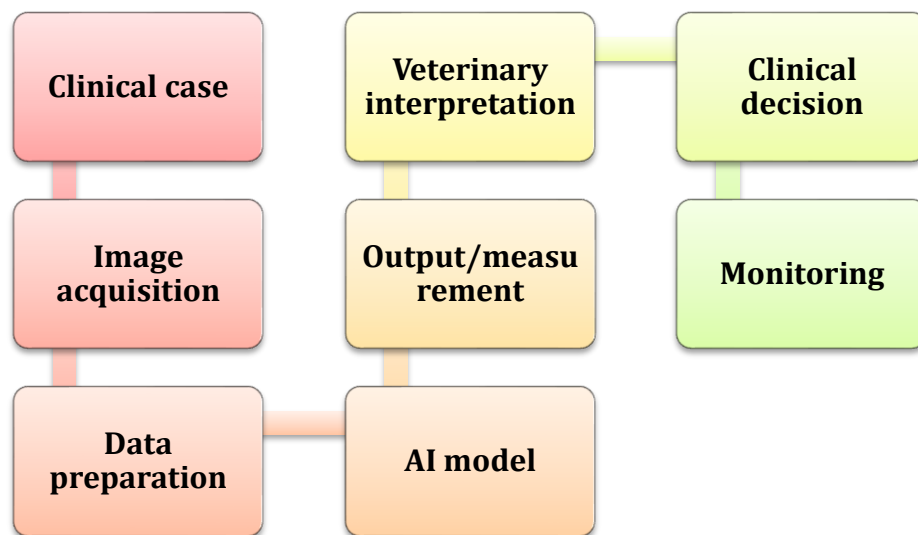


Figure 1: Conceptual workflow for veterinary diagnostic-imaging AI

3. Applications of AI in Veterinary Radiography

Radiography is one of the most established diagnostic imaging modalities in veterinary medicine and has generated a substantial proportion of veterinary AI research. Deep-learning systems have

been investigated for recognition and classification of thoracic abnormalities, cardiac enlargement and other radiographic findings. [5–7]

3.1 Thoracic Radiography

Thoracic radiograph interpretation can involve simultaneous assessment of the pulmonary parenchyma, airway, cardiovascular structures, pleural space and osseous structures. AI models have been developed to classify multiple thoracic findings. One published study used more than 22,000 canine and feline thoracic radiographs to train deep-learning models for 15 primary thoracic lesion categories. In an exploratory comparison, the best-performing network showed a lower overall error rate than the veterinarian observers in that study; the model was designed as a detection aid rather than an autonomous diagnostic system. [6]

3.2 Cardiac Assessment

Cardiac enlargement is commonly assessed on canine thoracic radiographs using measurements such as vertebral heart size (VHS). AI-assisted systems have been investigated for automated landmark detection and VHS estimation. Recent work has also explored automatic cardiothoracic ratio and VHS measurement in dogs and cats, aiming to improve speed and reproducibility. [8,9]

3.3 Musculoskeletal Radiography

AI may assist with detection and localization of fractures and other skeletal abnormalities. A 2026 study described a deep-learning approach for canine long-bone fracture classification using conventional radiographs, with automated segmentation incorporated to localize fracture regions. Such findings are promising but should be interpreted in the context of dataset composition, external validation and clinical workflow before generalization to routine practice. [11]

3.4 Dental Radiography

AI has also been evaluated in veterinary dental radiography. A 2024 study assessed a commercially available AI-based program for detection of common dental radiographic pathologies in dogs and cats, illustrating how specialized imaging tasks may become candidates for computer-assisted assessment. [12]

Table 1: Examples of AI applications in veterinary radiography

Imaging Area	Potential AI Task	Clinical Relevance
Thoracic radiography	Lesion classification/detection	Supports screening and prioritization
Cardiac radiography	VHS/CTR estimation	Objective quantitative measurement
Orthopaedic radiography	Fracture detection/segmentation	Supports localization and assessment
Dental radiography	Pathology detection	Assists systematic review

4.AI in Computed Tomography

CT provides cross-sectional and three-dimensional information and is particularly valuable for complex anatomical regions, trauma, orthopaedic disorders, skull and spine evaluation, and tumour assessment. AI can potentially reduce the time required for repetitive image-processing tasks and support quantitative analysis.

4.1 Automated Segmentation

Segmentation refers to delineating an anatomical structure, lesion or tissue from surrounding structures. Automated segmentation may be useful for bones, organs, tumours and other regions. In surgical patients, accurate segmentation can provide a basis for three-dimensional visualization and preoperative planning.

4.2 Lesion Detection and Quantification

AI systems can be trained to identify suspicious regions and estimate lesion dimensions or volume. In oncology, quantitative measurements may support longitudinal assessment. However, performance should be evaluated against appropriate expert reference standards and clinically relevant endpoints.

4.3 Three-Dimensional Surgical Planning

The combination of CT data, segmentation and three-dimensional reconstruction has potential value in veterinary surgery. Possible applications include complex fracture planning, assessment of anatomical relationships around masses, and planning of patient-specific approaches. AI may assist with the image-processing component, while surgical planning remains a clinical task.

5. AI in Magnetic Resonance Imaging

MRI produces multiple sequences and high-dimensional datasets, making interpretation and segmentation computationally demanding. Veterinary AI research has increasingly investigated automated segmentation and classification, particularly in canine imaging. A recent narrative review reported notable use of AI for automated MRI segmentation and described expanding applications across veterinary imaging modalities. [5]

5.1 Automated Segmentation

Segmentation can delineate the brain, spinal cord, lesions or other anatomical structures. Automated methods may reduce repetitive manual work and create quantitative measures such as volume.

5.2 Neurological Imaging

AI-assisted MRI analysis may support classification or localization of neurological abnormalities. Because breed, age, body size, imaging protocol and disease spectrum can vary considerably, models require broad datasets and careful external validation.

6. AI in Ultrasonography

Ultrasonography is widely used in veterinary practice because it is non-invasive, does not use ionizing radiation and provides real-time information. However, image appearance is strongly influenced by operator technique, transducer position, patient movement and machine settings. These factors make robust AI development challenging.

Potential applications include organ recognition, lesion classification, tissue characterization, automated measurement and image-quality assessment. Dynamic ultrasound may eventually benefit from AI systems that analyse sequences rather than single still images. Nevertheless, training datasets must reflect the range of operators, machines and patient presentations encountered in clinical practice.

7. Radiomics and Quantitative Imaging

Radiomics is the extraction of quantitative features from medical images. Features may describe intensity, texture, shape and spatial relationships. A typical radiomics workflow consists of image acquisition, pre-processing, segmentation, feature extraction and statistical or AI-based analysis. [13]

In veterinary medicine, radiomics may complement conventional visual interpretation by converting image information into measurable variables. Potential applications include characterization of masses, disease classification, treatment monitoring and prognostic research. However, radiomic features can be sensitive to acquisition parameters, reconstruction settings, segmentation choices and preprocessing. Standardization is therefore important.

8. Relevance to Veterinary Surgery

For a Veterinary Surgery and Radiology programme, the most relevant value of AI lies not only in image interpretation but also in quantitative assessment and planning. AI-assisted segmentation of CT or MRI datasets may facilitate three-dimensional anatomical reconstruction. This can be useful when a surgeon needs to understand the relationship between a lesion and adjacent bones, vessels or organs.

8.1 Fracture and Orthopaedic Cases

Automated fracture detection and segmentation may help identify abnormalities on radiographs and localize regions for closer review. In complex orthopaedic cases, CT-based three-dimensional reconstruction may support surgical planning.

8.2 Oncology and Surgical Planning

Tumour segmentation and volumetric measurement may assist in documenting lesion extent and monitoring changes. AI-assisted segmentation may also help create patient-specific three-dimensional representations, although surgical decisions require integration of imaging with clinical examination, pathology and other diagnostic information.

8.3 Image-Guided Procedures

Future systems may integrate real-time imaging with computer vision to support image-guided interventions. Such applications remain an emerging area and require rigorous evaluation of safety, accuracy and clinical benefit.

9. Advantages of AI in Veterinary Diagnostic Imaging

- Automation of repetitive image-analysis tasks.
- Rapid processing of large image datasets.
- Objective and reproducible quantitative measurements.
- Potential support for detection of subtle or suspicious regions.
- Assistance with case prioritization and workflow management.
- Potential support for three-dimensional reconstruction and surgical planning.
- Generation of structured measurements or preliminary reports for clinician review.
- Research opportunities through large-scale quantitative image analysis.

10. Limitations and Challenges

10.1 Dataset Size and Diversity

Veterinary datasets are often smaller than human medical datasets and may be concentrated in particular breeds, institutions or equipment types. A model trained on a narrow population may not generalize well to other clinical environments. Recent reviews continue to identify dataset diversity and external validation as important barriers. [1,4,5]

10.2 Image Acquisition Variability

Differences in positioning, exposure, reconstruction, field of view, equipment and operator technique can alter image appearance. These variations must be represented in development and validation datasets.

10.3 Ground-Truth Labels

Supervised learning requires reference labels. In imaging, these may be radiologist reports, consensus annotations, pathology results, follow-up findings or quantitative measurements. The quality of the reference standard directly influences the model.

10.4 Class Imbalance

Some diseases are common while others are rare. A dataset containing mostly normal or common cases may produce a model that performs well on the dominant classes but poorly on rare clinically important abnormalities.

10.5 Generalizability and External Validation

Internal test performance does not guarantee clinical performance. Multi-centre, multi-device and temporally separated validation can provide stronger evidence of generalizability.

10.6 Explainability and Uncertainty

Deep-learning models may be difficult to interpret. Clinicians need to understand not only the predicted class but also the limitations and uncertainty of the system. Visual explanations may help, but they should not be treated as proof that the model used clinically meaningful features.

Table 2: Major challenges in veterinary imaging AI and possible approaches

Challenge	Why It Matters	Possible Approach
Small datasets	Risk of overfitting	Multi-centre data, transfer learning, augmentation
Breed/species variation	Different anatomy and disease prevalence	Diverse datasets and stratified evaluation
Poor labels	Model learns incorrect targets	Expert consensus and clinically meaningful reference standards
Equipment/protocol variation	Distribution shift	External validation across devices/sites
Rare diseases	Low sensitivity for uncommon classes	Targeted sampling and appropriate metrics
Black-box behaviour	Difficult clinical interpretation	Explainability tools and uncertainty reporting

11. Ethical, Legal and Professional Considerations

AI implementation in veterinary diagnostic imaging requires attention to patient data security, transparency, accountability and post-implementation monitoring. The 2025 position statement from the American College of Veterinary Radiology and European College of Veterinary Diagnostic Imaging recommends good machine-learning practices, including transparency, error reporting, secure handling of patient data and involvement of clinical experts throughout development. It also emphasizes maintaining a veterinarian in the loop. [10]

A 2026 clinical review similarly describes AI-assisted radiograph interpretation as a decision-support function rather than a replacement for clinical judgment, with veterinarians retaining professional responsibility for diagnosis and treatment decisions. [14]

11.1 Data Privacy

Imaging datasets should be stored and transferred securely. When datasets are used for research or commercial development, appropriate institutional and legal requirements for data governance should be followed.

11.2 Accountability

The person responsible for the clinical case must understand the role and limitations of the AI tool. AI output should be documented appropriately when it materially influences clinical workflow.

11.3 Continuous Monitoring

Performance can change when equipment, protocols, patient populations or disease prevalence change. Post-implementation monitoring and error reporting are therefore important components of responsible deployment.

12. Future Perspectives

The next phase of veterinary imaging AI is likely to move beyond isolated image-classification tasks toward integrated systems that combine multiple imaging modalities and clinical information.

- Multimodal AI combining radiographs, CT, MRI, ultrasound and clinical information.
- Automated segmentation and quantitative measurement.
- AI-assisted structured reporting with mandatory clinician review.
- Integration with PACS and electronic medical records.
- Three-dimensional surgical planning and patient-specific anatomical models.
- Radiomics combined with clinical and laboratory variables.
- Real-time computer vision for image-guided procedures.
- Federated or privacy-preserving approaches for multi-centre model development.
- Prospective clinical trials evaluating real-world clinical benefit.

Recent literature indicates that veterinary AI imaging research is expanding rapidly, particularly in dogs and in radiography and MRI-related applications. [5] At the same time, current work continues to emphasize the need for external validation, diversity of datasets, interpretability and appropriate clinical oversight. [4,5,14]

13. Practical Recommendations for Developing a Veterinary Imaging AI Study

For researchers in Veterinary Surgery and Radiology, the following workflow can improve the scientific quality of an AI imaging project:

- Define one clinically meaningful question rather than attempting to solve multiple unrelated tasks.
- Specify the target species, breed range, age group, imaging modality and disease spectrum.
- Create clear inclusion and exclusion criteria before model training.
- Use high-quality reference standards and, where appropriate, expert consensus.
- Separate training, validation and test data at the patient level to avoid data leakage.
- Where feasible, reserve an external dataset from another institution or time period.
- Report clinically meaningful performance measures such as sensitivity, specificity, precision, F1 score and area under the ROC curve, as appropriate to the task.
- Evaluate performance by relevant subgroups such as species, breed, projection or equipment where sample size permits.

- Document preprocessing, model architecture and training procedures sufficiently for reproducibility.
- Include veterinary radiologists or relevant clinical specialists throughout study design and evaluation.
- Assess whether the model improves clinical workflow or patient care rather than relying only on technical accuracy.
- Plan for monitoring and reporting errors after implementation.

Conclusion

Artificial intelligence represents an important emerging technology in veterinary diagnostic imaging. Applications have expanded from experimental image classification toward more clinically oriented tasks such as thoracic lesion detection, cardiac measurement, fracture assessment, segmentation and quantitative analysis. Current literature demonstrates potential across radiography, CT, MRI and ultrasound, with dogs being a major focus of published work. [4,5]

For Veterinary Surgery and Radiology, AI has particular relevance to quantitative imaging, orthopaedic assessment, oncological imaging, three-dimensional reconstruction and surgical planning. Nevertheless, promising technical performance should not be equated automatically with clinical effectiveness. Dataset quality, generalizability, external validation, transparency, data governance and clinical oversight remain essential.

The most appropriate future model is therefore collaborative: AI performs computationally intensive or repetitive tasks, while veterinary professionals provide clinical context, interpretation and final decision-making. With rigorous validation and responsible governance, AI may become an increasingly useful component of veterinary imaging and surgical workflows.

Table 3: Summary of major veterinary imaging modalities and potential AI roles

Modality	Example AI Task	Potential Veterinary Use	Key Limitation
Radiography	Classification/detection / measurement	Thoracic, cardiac, orthopaedic screening	Positioning and dataset variation
Ultrasound	Classification / measurement	Organ and lesion assessment	Operator dependence
CT	Segmentation / 3D reconstruction	Complex anatomy, oncology, surgery	Large datasets and protocol variation
MRI	Segmentation / classification	Neurology and soft tissue	Multiple sequences and limited datasets

References

1. Hennessey, E., DiFazio, M., Hennessey, R., & Cassel, N. (2022). Artificial intelligence in veterinary diagnostic imaging: A literature review. *Veterinary Radiology & Ultrasound*, 63(Suppl. 1), 851–870. <https://doi.org/10.1111/vru.13163>
2. Hespel, A. M., Zhang, Y., & Basran, P. S. (2023). Artificial intelligence 101 for veterinary diagnostic imaging. *Veterinary Radiology & Ultrasound*. <https://doi.org/10.1111/vru.13160>
3. Banzato, T., Wodzinski, M., Burti, S., Longhin Osti, V., Rossoni, V., Atzori, M., & Zotti, A. (2021). Automatic classification of canine thoracic radiographs using deep learning. *Scientific Reports*, 11, 3964. <https://doi.org/10.1038/s41598-021-83515-3>
4. Burti, S., Banzato, T., Coghlan, S., Wodzinski, M., Bendazzoli, M., & Zotti, A. (2024). Artificial intelligence in veterinary diagnostic imaging: Perspectives and limitations. *Research in Veterinary Science*, 175, 105317. <https://doi.org/10.1016/j.rvsc.2024.105317>
5. Barillaro, G., Minniti, S., Mezzatesta, S., Macrì, F., Torino, C., & Vilasi, A. D. (2026). Artificial intelligence in veterinary diagnostic imaging: A narrative review of the state of the art and its applications. *The Veterinary Journal*, 318, 106762. <https://doi.org/10.1016/j.tvjl.2026.106762>
6. Boissady, E., de La Comble, A., Zhu, X., & Hespel, A. M. (2020). Artificial intelligence evaluating primary thoracic lesions has an overall lower error rate compared to veterinarians or veterinarians in conjunction with the artificial intelligence. *Veterinary Radiology & Ultrasound*. <https://doi.org/10.1111/vru.12912>
7. Banzato, T., Wodzinski, M., Burti, S., *et al.* (2021). Automatic classification of canine thoracic radiographs using deep learning. *Scientific Reports*, 11, 3964. <https://doi.org/10.1038/s41598-021-83515-3>
8. Mekonnen, *et al.* (2025). Deep learning framework for vertebral heart size and cardiothoracic ratio estimation in dogs and cats using thoracic radiographs. *Scientific Reports*. Study indexed in PubMed.
9. Automated vertebral heart size estimation from thoracic radiographs in dogs with AI-assisted clinical decision support. (2026).
10. Appleby, R. B., DiFazio, M., Cassel, N., Hennessey, R., & Basran, P. S. (2025). American College of Veterinary Radiology and European College of Veterinary Diagnostic Imaging position statement on artificial intelligence. *Journal of the American Veterinary Medical Association*, 263(6), 773–776. <https://doi.org/10.2460/javma.25.01.0027>
11. Saber, A. S., Selim, I., Askr, H., *et al.* (2026). Veterinary fracture diagnosis: A deep learning model for dogs long bone fractures. *Scientific Reports*, 16, 15098. <https://doi.org/10.1038/s41598-026-50387-4>

12. Nyquist, M. L., Fink, L. A., Mauldin, G., & Coffman, C. R. (2024). Evaluation of a novel veterinary dental radiography artificial intelligence software program. *Journal of Veterinary Dentistry*. <https://doi.org/10.1177/08987564231221071>
13. Bouhali, O., Bensmail, H., Sheharyar, A., David, F., & Johnson, J. P. (2022). A review of radiomics and artificial intelligence and their application in veterinary diagnostic imaging. *Veterinary Sciences*, 9(11), 620. <https://doi.org/10.3390/vetsci9110620>
14. Basran, P. S., Appleby, R., & Porter, I. (2026). Artificial intelligence-assisted interpretation of veterinary radiographs: Opportunities, risks, and best practices for clinicians. *Veterinary Clinics of North America: Small Animal Practice*, 56(5), 1105–1115. <https://doi.org/10.1016/j.cvsm.2026.03.018>
15. Vickram, A. S., *et al.* (2025). AI-powered techniques in anatomical imaging: Impacts on veterinary diagnostics and surgery. *Annals of Anatomy*. Study indexed in PubMed.
16. Review of applications of deep learning in veterinary diagnostics and animal health. (2025). *Frontiers in Veterinary Science*.

TALENT ACQUISITION IN THE ERA OF ARTIFICIAL INTELLIGENCE: A REVIEW OF CURRENT TRENDS AND CHALLENGES

Monika Chauhan and Nisha Agarwal

Department of Management Studies,

School of Business Management, IFTM University, Moradabad, India

Corresponding author E-mail: monika.chauhan@iftmuniversity.ac.in,
nishaagarwal@iftmuniversity.ac.in

Abstract

Artificial Intelligence (AI) is revolutionizing talent acquisition through candidate sourcing, resume screening, communication, interview support, and workforce analytics. This chapter examines the state of AI-powered recruitment solutions and highlights their advantages for recruitment, efficiency, scalability, candidate experience, and decision-making. It also addresses some of the major issues like algorithmic bias, transparency, ethical issues, candidate trust, etc. The chapter emphasizes the role of human oversight, the need for responsible government and management of AI and the fairness in the recruitment process. Lastly, it explores new avenues for future research, such as generative and agentic AI, bias mitigation, regulatory changes, and human-AI collaboration in talent acquisition.

Keywords: Artificial Intelligence, Recruitment, Talent Acquisition, Algorithmic Bias and Fairness, Human-AI Collaboration, Workforce Analytics, Generative and Agentic AI.

1. Introduction

Sourcing candidates, reviewing applications, scheduling interviews, and making quick calls as deadlines approach - recruitment has always been a high-effort, high-pressure job function. The last decade introduced AI to these processes thanks to faster, more sophisticated machine learning and natural language processing.

Now, a combination of surveyed organizations and organizations willing to go on the record attest to there being a larger, exponentially growing share of companies with AI at a minimum in some stage of their recruitment process. This paper looks at how organizations are using it, the positive impacts it's had, the drawbacks and the ethical debates surrounding its use, and how laws are changing to address it.

2. Evolution of AI in Recruitment

Ancient implementation of technology in recruitment was mainly creating an application tracking system (ATS), which was used solely to store resumes and perform keyword-based searches. The first major milestone affecting the ability to process resumes was the emergence of

NLP for natural language processing, and the next was a machine learning model, which could rank candidates based on data patterns derived from past hires. [1]; [2]

Recent years, the advent of Generative AI and LLM has broadened the possibilities. You're not merely matching keywords— modern systems can read through unstructured job postings and resumes, produce personalized outreach messages, do initial conversational filters and summarize interview calls. The latest trend is the use of agentic AI, where a completely autonomous system performs various parts of the recruiting process—from scheduling, follow-up email, initial candidate screening, and more—with very little manual oversight required until a decision is made. [3] The adoption of such agents by enterprises has surged rapidly – per one pulse survey in the industry, enterprises with AI agents in production more than tripled over a single quarter in 2025, and a global survey of talent leaders revealed that over half will introduce autonomous recruiting agents in the year. [3]

3. Current Applications of AI across the Recruitment Lifecycle

3.1 Sourcing and Talent Discovery: AI-powered sourcing tools comb through job boards, professional networks, and internal talent pools to find the passive candidates that could be a fit for a role, not in the process of applying. Such systems are becoming more common that use semantic matching instead of simple keyword searching; that means they can identify that the experiences that a candidate is looking for are relevant even if not stated exactly as the job posting. [4]; [5]; [6].

3.2 Resume Screening and Candidate Matching: Automated screening continues to be the most common AI use case in recruitment. These AI-powered tools scan resumes, extracting key details like work experience, skills, and education, before prioritizing or scoring them according to a job's requirements. A significant portion of companies implementing AI in recruiting apply the AI technology to this initial screening of resumes, as it's during this stage that the volume issue arises the most. [7], [3], [8]; [9], [10].

3.3 Candidate Communication and Engagement: A significant proportion of the candidate interaction that occurs is now managed by a chatbot and a conversational assistant, which can answer very common queries, track application status and arrange interview date and times on various calendars. Tracking the scheduling of interviews can save even a huge amount of time for the back-and-forth communication when using automated scheduling. [11] According to candidate surveys, job seekers are most open to AI at the beginning of the funnel: Popular methods for interactions with candidates for initial inquiries and scheduling include using a chatbot, with the majority of still wanting a human touch at the end. [4]

3.4. Interview Support and Assessment: The number of organisations employing AI for interview support and/or part of the interview is modest but slightly increasing and includes companies employing AI to create structured and job-specific interview questions, transcribe

conversations, as well as generate summary documents for hiring panels. Some tools try to analyse the language behaviour for indicators of communication-related skills, but this is one of the most debated areas, because of doubts in terms of validity and fairness. In terms of candidate-facing surveys, it is found that only around a quarter of job candidates believe that AI is fair towards them [4] [12], even as a majority state that they have already experienced at least one AI-based interview. [12],[9],[13]

3.5 Workforce and Predictive Analytics: Outside of the hiring event, AI is used to predict hiring requirements, estimate time-to-fill for roles and in more sophisticated setups, to identify signs of employee attrition at a very early stage. There's also a rise in the use of AI skills-matching-driven internal talent marketplaces to support organizations in filling roles with internal mobility over external recruitment. [3],[14], [15]

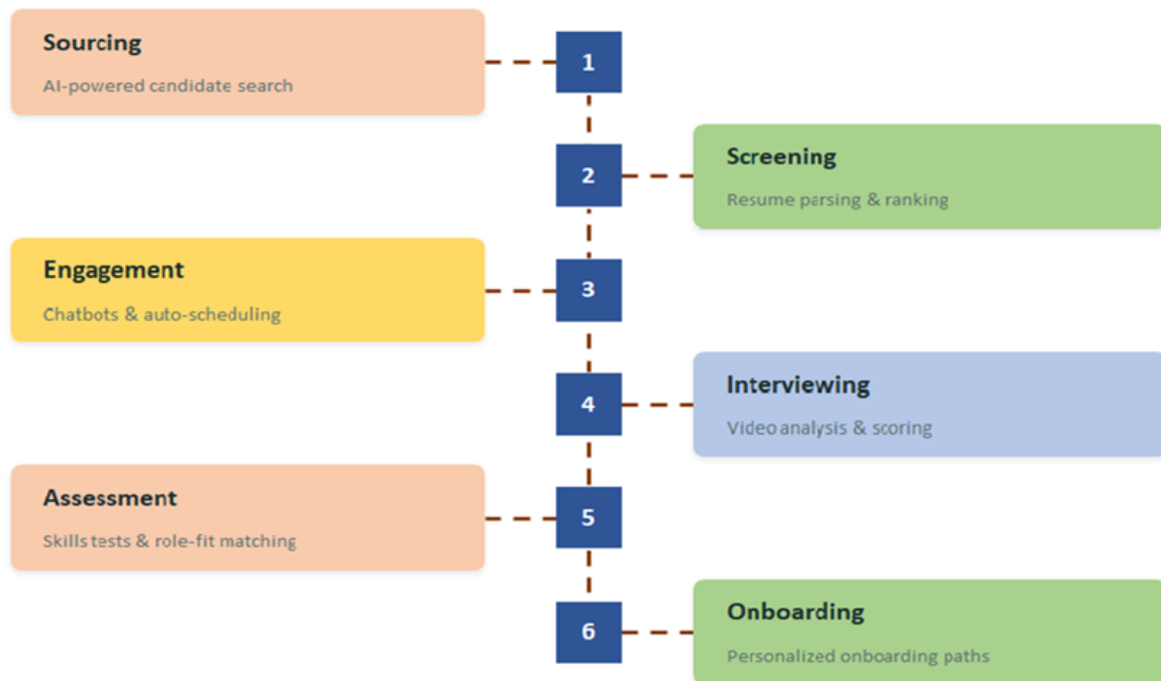


Figure 1: AI-Powered Recruitment Process

4. The Current State of Adoption in Organization

The trend is that artificial intelligence in recruitment is no longer in the experimental phase, but making its way into organizations. Empirical Research indicated that technology readiness, organizational capability, perceived usefulness, management support, regulations and human factors affect the adoption process. [16] There's also evidence that AI is replacing human talent-acquisition processes, such as resume screening, candidate assessment, forecasting, and analysis. It is being adopted, but implementation varies from organization to organization. [17],[18],[19] Therefore, percentages of industry surveys do not adequately capture the current state of adoption, with data needed to learn the adoption drivers; Table1 shows organizational readiness, actual adoption, and implementation practices.

Table 1: Empirical evidence on AI adoption in recruitment and talent acquisition.

Adoption Indicator	Research evidence	Interpretation	Citation
AI adoption in employee recruitment	297 Chinese companies surveyed	Perceived AI complexity constraints adoption, while technology competence and regulatory support encourage adoption	Pan <i>et al.</i> (2022)[20]
AI adoption for talent acquisition	562 HR and talent acquisition managers	The factors that support adoption of cost effectiveness and relative advantage, top management support, and HR readiness have positive influences, while security/privacy concerns have a negative influence. Cost effectiveness and relative advantage, top management support, and HR readiness are positive influences while security/privacy concerns are negative.	Pillai Rajasshrie and Sivanthanu Brijesh (2020)[21]
Effective AI adoption in multinational talent acquisition	19 cases across seven industries	Adoption requires explicit motives, barriers, steps and critical success factors.	Marvi <i>et al.</i> (2024)[19]
AI adoption and harmful practices in talent acquisition	42 Expert interviews + 303 survey participants	The adoption of AI and its impacts are influenced by both technological and organizational/economical and individual conditions.	Roppelt <i>et al.</i> . (2025)[22]
AI adoption in recruitment among HR Professionals	238 recruiters surveyed	Prior frequency of use and performance expectancy both play a role in intention to use AI; efficiency and time savings are seen as significant benefits.	Ore & Sposato (2025)[23]
AI-based recruitment adoption in Thailand	HR and recruitment professionals	Aligned to the UTAUT approach for the adoption intention, the supporting factors of the performance expectation, effort expectation, social influence and facilitating conditions are being explored.	Islam <i>et al.</i> . (2024) [24]
AI –based recruitment readiness in Indian Organization	126 Indian companies surveyed	Delivers in-depth analysis of organization preparedness and readiness to implement an AI recruitment tool in various industries.	Gadesha & Joshi (2025)[25]

AI adoption in Indian IT-sector HRM	Empirical structural model study	Factors that influence the adoption of AI within organizations are explored, including the issue of competition, security/privacy concerns and technological readiness.	Durai & Anbu (2026)[26]
AI use across recruitment activities	Systematic review of 49 peer-reviewed studies	he research shows expanding applications discovered in screening and assessment, analytics and other recruitment activities along with privacy, interference with bias and transparency issues.	Dadaboyev <i>et al.</i> (2025)[17]
AI adoption in Indian IT recruitment	14 HR and IT professionals from four Indian IT companies	New opportunities for automated resume screening, chatbots, and predictive analytics, and challenges of integration and resistance to change are all becoming more prevalent.	Itam <i>et al.</i> (2026)[27]

Benefit of Organization for Adoption of AI Recruitment

- **Efficiency gains:** A significant number of firms adopting AI for recruitment report tangible efficiency gains, with projected Time-to-Hire reductions spanning from approximately 25% to 70% contingent on the level of integration and usage of the technology in their recruitment practices.[3],[14],[18]
- **Cost savings:** significant cost savings were attributed to most of the advertisers who say that they reduced their non-manual effort in screening prospective hires, as well as to their improved scheduling times.[14]
- **Consistency in evaluation:** it can structure and facilitate the evaluation process by establishing criteria and eliminating the use of ad hoc human evaluation, which can at times be communicated differently across the board. [23]
- **Enhanced candidate experience:** studies have shown that most candidates have a more positive experience when they receive timely updates and receive a speedy response during the hiring process due to the AI chatbot . [4]
- **Scalability:** AI can support recruiting agencies with a much larger number of applications without any proportional growth in their size, as large-scale or seasonal recruitment is common.[28]
- **Reported Business:** If an organization has implemented AI in HR, a majority of surveyed HR managers describe efficiency gains in their work, and improvements in the quality of the work .[28]

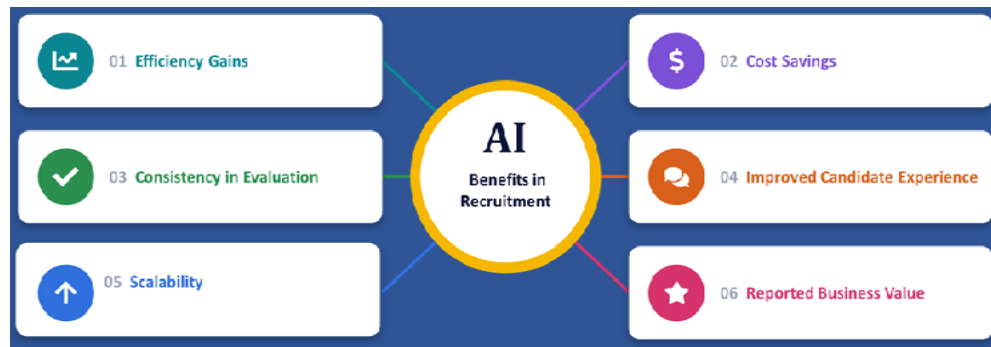


Figure 2: Benefits of AI in Recruitment [3],[4],[14],[28]

5. Challenges and Risks

5.1 Algorithmic Bias

One of the most constant fears around AI for recruitment is the possibility for it to perpetuate or reinforce the prejudices within the Victorian hiring information. The more complex systems that utilize machine learning algorithms will learn from past choices, meaning if they are trained exclusively from a less diverse workforce, these systems can subconsciously discriminate against people from underrepresented communities. This worry has a real-world resonance that preexisted generative AI: A study of a field experiment revealed that fake resumes with white-sounding names got far more callbacks than fake resumes with black-sounding names. [29] One popular real world example has been an internal real-time, AI recruiting system that one large tech firm ended up losing because it had become self-innulating data from the last 10 years of historically male dominated hiring data. [30] Similarly, more recent controlled studies of the language-model-based resume screening indicate similarly large disparities: one recent academic study of over 500 resumes across nine occupations found that the models favored white-associated names by around 85 percent, with a lower percentage favoring black-associated names of around 9 percent, and with similar gaps between male and female resumes. [9],[29], [31]

5.2 Transparency and Explainability

The majority of AI hiring tools work like a black box; they do not explain their decision-making process, meaning that customers have few indications of why a specific candidate may have been preferred over another. This opaqueness poses issues for candidates to be provided with feedback and for organisations to run their own fairness audit.[32],[33],[34]

5.3 Over-Reliance and Loss of Human Judgment

The danger is that recruiters rely too much on AI-produced scores and do not consider them an absolute but one of several factors. This is especially problematic for jobs that aren't a perfect fit with the skills they've previously seen in their historical hiring record, such as a cultural fit or a career progression that's not strictly linear. Hopefully, most of the recruiters are not just giving it all to the AI tool: one industry survey suggests the vast majority of respondents wished to have

"supervisory authority" over whatever recommendations the tool provided rather than hand everything over [34],[35],[36]

5.4 Candidate Trust

But the confidence in hiring decisions made by AI is still mixed and varied. The statistics from the survey also indicate an increase in confidence in AI hiring tools from approximately 37 percent in 2024 to just over half of respondents in 2025, with small gains in the percentage of people who feel AI is making hiring more equitable by reducing racial and ethnic bias.[36] Other 2026 surveys present a more pessimistic outlook: A multi-country applicants survey showed that only around a quarter of the job seekers believe AI can fairly evaluate them, although nearly two-thirds of applicants surveyed in the U.S. stated they have faced an AI-supported interview within the past six months. [36] Another US worker survey likewise revealed that, despite the growing extent to which many workers today experience some level of AI within their hiring or workplace processes, age and job type created a wide gap between workers that set the tone for their trust and comfort levels.[37] Nevertheless, a significant percentage of candidates are still wary, especially about automated interviewing assessment.[35],[9],[38]

6. Implications for Organizations

When companies are thinking about adopting AI in recruitment or looking into expanding its reach, evidence suggests certain implications. First, AI provides the greatest value at the beginning of the hiring journey, where volumes are the highest, and manual review costs the most—sourcing and initial screening. Second, the risk of bias and legal liability increases the further it can get into an employer's hiring process, so having human involvement at the point of deciding is especially critical — which aligns with the preference of recruiters who say that they want to have final say in making the hiring decision. [38] Third, there is a growing body of jurisdictions where it is no longer merely a best practice to be transparent about the use of AI with candidates, but a law of the land. Third, while there are specific dislike deadlines still shifting, it is a growing body of jurisdictions where parties expecting to use AI with candidates are expected, if not required, to be when and how it is used. [39], [2],[18] Successful AI adopters generally view AI as a decision-support tool that identifies the information and recommendations that should be passed on to a human recruiter, not reaching conclusions on their own and making the hiring decision. This blend does seem to take advantage of many of the efficiencies of AI, but it also maintains accountability and limits legal and reputational liability. [3]

7. Future Scope of AI in Recruitment

The future research and practice directions for AI in recruitment highlight the need for further investigation and development. Firstly, more embedded agentic systems will be likely in the future, continuing the conceptual exploration lead by [3] in their discussion of the impact of

autonomous intelligent systems on organisational recruitment and selection. Second, mitigation of the different forms of bias, such as that of hiring, will continue to be one of the important research fields as indicated by [30], [9] who propose the refining of the frameworks of behavioural and related research in this regard and Seppälä and Małecka (2024) [40] who suggest that work is required on the avenues of mitigation of discriminatory outcomes towards ensuring that the applications of AI are without bias.

Third, research around the candidate experience will likely continue to grow, as Bhattacharya, Gupta, and Horodyski (2023)[41] note, because of the lack of knowledge regarding the applicants' trust and perception of AI-assisted hiring decisions. Fourth, there is a need for research on both the longitudinal and cross-industry levels to validate claims of efficiency by [4] and [42] for the actual effects of AI on selection.

Looking forward, the landscape can be brightened with the shift from adoption studies focused on description to adoption studies that focus on evaluation and ethics such as regulatory compliance, fairness auditing, models of human-AI interaction, and even the longer-term impacts on organizational diversity and candidate trust.

Conclusion

From a relatively new technology to a growing cornerstone of modern recruitment, AI is becoming prevalent in hiring. What used to be a relatively new technology is now a very significant aspect of modern recruitment and time using AI is now a much-anticipated necessity. Its use can streamline workflow, facilitate resume reading and writing, help for interviews and insights on the workforce. Meanwhile, its increasing popularisation poses an issue of bias, transparency, accountability, and trust of candidates. To harness the true potential of AI, it is crucial to strike a balance and ensure that technology complements, but doesn't supplant, human decision-making. Adopting AI with proper human oversight, process transparency, ethical measures, and responsible governance practices can help organizations derive more value from this technology. Finally, the future of AI-powered recruitment relies on human intention and technologic abilities to put together effective, fair, transparent and trustworthy recruiting procedures.

References

1. Herold, M., & Roedenbeck. (2025). Cultural values and the P-O fit: Comparative NLP analysis of German online job advertisements. *Evidence-Based HRM*, 13(2).
2. Williams, P., & McDonald, P. (2026). Artificial intelligence in recruitment and selection: A conceptual model and research agenda. *Journal of Information, Communication and Ethics in Society*, 32(11). <https://doi.org/10.1108/JICES-10-2025-0269>
3. Bujold, A., et al. (2024). Responsible artificial intelligence in human resources management: A review of the empirical literature. *AI and Ethics*, 4(4), 1185–1200. <https://doi.org/10.1007/s43681-023-00325-1>

4. Van Esch, P., Black, J. S., & Ferolie, J. (2019). Computers in Human Behavior marketing AI recruitment: The next phase in job application and selection. *Computers in Human Behavior*, 90(September 2018), 215–222. <https://doi.org/10.1016/j.chb.2018.09.009>
5. Jeske, D. (2019). Social media screening and content effects: Implications for job applicant reactions. *International Journal of Manpower*, 40(1).
6. Kong, Y. (2023). Tools, potential, and pitfalls of social media screening: Social profiling in the era of AI-assisted recruiting. *Journal of Business and Technical Communication*, 38(1).
7. Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, 178.
8. Raghavan, M., Barocas, S., & Kleinberg, J. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*.
9. Köchling, A. (2020). Discriminated by an algorithm: A systematic review of discrimination and fairness by algorithmic decision-making in the context of HR recruitment and HR development. *Springer Nature Link*, 13.
10. Fabris, A., Baranowska, N., Dennis, M. J., Graus, D., Hacker, P., Saldivar, J., & Zuiderveen Borgesius, F. (2025). Fairness and bias in algorithmic hiring: A multidisciplinary survey. *ACM Transactions on Intelligent Systems and Technology*, 16(1).
11. Kshetri, N. (2021). Evolving uses of artificial intelligence in human resource management in emerging economies in the global South: Some preliminary evidence. *Management Research Review*, 44(7).
12. Suen, H.-Y., & Lu, C.-L. (2020). Intelligent video interview agent used to predict communication skill and perceived personality traits. *Springer Nature Link*, 10.
13. Acikgoz, Y., Davison, K. H., & Compagnone, M. (2020). Justice perceptions of artificial intelligence in selection. *International Journal of Selection and Assessment*, 28(4).
14. Tambe, P., Cappelli, P., & Yakubovich, V. (2019). Artificial intelligence in human resources management: Challenges and a path forward. *California Management Review*, 61(4).
15. Strohmeier, S., Becker, M., & Scheer-Weller, E. (2026). Beyond aversion – principles of appropriate algorithmic decision-making in human resource management. *Expert Systems with Applications*, 296(PA), 128954. <https://doi.org/10.1016/j.eswa.2025.128954>
16. Gül, B., & Soyer, A. (2026). AI-based recruitment: An integrative framework for human resources professionals' adoption. *MDPI Open Access Journals*, 14(6).
17. Dadaboyev, S. M. U., Abdullayeva, J., & Abbosova, N. (2025). Role of artificial intelligence in employee recruitment: Systematic review and future research directions. *Springer Nature Link*, 3.

18. Dutta, D. (2025). Transforming recruitment and selection practices in organizations through discriminative and generative AI adoption: A structuration lens. *Human Resource Management*, 65(1).
19. Marvi, R., Schuster, A., & AmirDadbar. (2024). Artificial intelligence in talent acquisition: A multiple case study on multi-national corporations. *Management Decision*, 62(10).
20. Pan, Y. (2022). The adoption of artificial intelligence in employee recruitment: The influence of contextual factors. *International Journal of Human Resource Management*, 33(6).
21. Pillai, R. (2020). Adoption of artificial intelligence (AI) for talent acquisition in IT/ITeS organizations. *Benchmarking: An International Journal*, 27(9).
22. Roppelt, J. S., Schuster, A., Kanbach, K. S., & Dominik, K. (2025). Towards effective adoption of artificial intelligence in talent acquisition: A mixed method study. *International Journal of Information Management*, 82.
23. Author, O. O. M. S. (30AD). Opportunities and risks of artificial intelligence in recruitment and selection. *International Journal of Organizational Analysis*, 6.
24. Islam, M., Rahman, M. M., & Taher, M. A. (2024). Using artificial intelligence for hiring talents in a moderated mechanism. *Springer Nature Link*, 10(13).
25. Gadesha, V., & Joshi, K. (2025). Exploring the potential of an AI-based recruitment system for Industry 4.0 landscape: A survey-based study. *Vikalpa*. <https://doi.org/10.1177/02560909251379212>
26. Durai, A. K., & Anbu, A. (2026). Identification of artificial intelligence adoption determinants affecting human resource management effectiveness in the Indian information technology sector. *Eastern-European Journal of Enterprise Technologies*, 2. <https://doi.org/10.15587/1729-4061.2026.357122>
27. Itam, U. J., Shukla, N., & Toni, M. (2026). AI-driven talent acquisition in Indian IT firms: A qualitative study of automation, impacts, and skill requirements. *Springer Nature Link*, 6. <https://doi.org/10.1007/s43546-026-01325-z>
28. Paramita, D., & Okwir, S. (2024). Artificial intelligence in talent acquisition: Exploring organisational and operational dimensions. *International Journal of Organizational Analysis*, 32(11). <https://doi.org/10.1108/IJOA-09-2023-3992>
29. Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *American Economic Review*, 94(4).
30. Raghavan, M., Barocas, S., Kleinberg, J., & Levy, K. (2020). Mitigating bias in algorithmic hiring: Evaluating claims and practices. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*.

31. Drage, E., & Mackereth, K. (2022). Does AI debias recruitment? Race, gender, and AI's "eradication of difference". *Philosophy & Technology*, 35, 1–25.
32. Lepri, B., Oliver, N., Letouzé, E., Pentland, A., & Vinck, P. (2018). Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*, 31.
33. Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference*.
34. Geden, J., & Andrews, M. (2021). Fair and interpretable algorithmic hiring using evolutionary many objective optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*.
35. Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1).
36. Pessach, D., & Shmueli, E. (2021). Improving fairness of artificial intelligence algorithms in privileged-group selection bias data settings. *Expert Systems with Applications*, 185.
37. Suen, H.-Y., Chen, M.-Y.-C., & Lu, S.-H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior*, 98.
38. Babae, A., & Shank, D. B. (2025). Trust and procedural justice mediate the effects of AI hiring on organizational attraction and intent to apply. *Discover Artificial Intelligence*, 5.
39. Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, 178.
40. Seppälä, P., & Małecka, M. (2024). AI and discriminative decisions in recruitment: Challenging the core assumptions. *Big Data & Society*, 11(1).
41. Horodyski, P. (2023). Applicants' perception of artificial intelligence in the recruitment process. *Computers in Human Behavior Reports*, 11.
42. Upadhyay, A. K. (2018). Applying artificial intelligence: Implications for recruitment. *Strategic HR Review*, 17(5).

AI-DRIVEN SMART SYSTEMS IN BUSINESS MANAGEMENT

Divyansh Mishra¹, Rajesh Kumar Mishra² and Rekha Agarwal³

¹XLRI – Xavier School of Management, Jamshedpur, Jharkhand 831001

²ICFRE-Tropical Forest Research Institute

(Ministry of Environment, Forests & Climate Change, Govt. of India)

P.O. RFRC, Mandla Road, Jabalpur, MP-482021, India

³Department of Physics, Government Science College, Jabalpur, MP, India- 482 001

Corresponding author E-mail: divyanshspps@gmail.com, rajeshkmishra20@gmail.com,
rajesh.mishra0701@gov.in, rekhasciencecollege@gmail.com

Abstract

Artificial intelligence (AI) has become inseparable from the broader computing and data infrastructure that enterprises rely on to sense, analyze, and act on information in real time. This chapter examines how AI-enabled smart systems -- built on cloud and edge computing infrastructure, large-scale data pipelines, and increasingly autonomous decision agents -- are reshaping core business management functions, including strategy, human resource management, marketing, operations, and finance. Drawing on peer-reviewed research, systematic reviews, and current industry survey evidence, the chapter develops an integrated view of the technical and organizational conditions under which AI-driven smart systems generate measurable business value. It examines the data and computing stack underlying enterprise AI, the organizational enablers and barriers to adoption, the ethical and governance dimensions of deploying autonomous or semi-autonomous systems in consequential business decisions, and the labor-market implications of this transition. The chapter concludes that computing infrastructure and data quality function as necessary but not sufficient conditions for AI-driven value creation: organizational redesign, governance maturity, and workforce capability-building remain the decisive factors distinguishing organizations that convert smart-systems investment into sustained competitive advantage from those that do not.

Keywords: Artificial Intelligence, Smart Systems, Cloud and Edge Computing, Data Governance, Business Management, Digital Transformation.

1. Introduction

Artificial intelligence is increasingly deployed not as a standalone analytical tool but as one layer within broader smart systems architecture -- an interconnected stack of sensors, data pipelines, cloud and edge computing resources, and decision agents that together allow organizations to sense conditions, process information, and act with a speed and scale unavailable to purely human-mediated processes. This chapter examines AI-driven smart systems specifically through a business management lens, asking how the underlying data and computing infrastructure

discussed elsewhere in this volume translates into organizational capability, and what conditions determine whether that capability converts into measurable business value.

The chapter's central argument is that computing infrastructure and data availability, while necessary preconditions for AI-driven smart systems, are not sufficient conditions for value creation. Global survey evidence indicates that AI usage in at least one business function rose from 78 percent of organizations in 2024 to 91 percent by 2026, with generative AI usage nearly doubling from 37 percent in 2023 to 72 percent by 2025, as shown in Figure 1 (McKinsey & Company, 2026a). Yet the same evidence indicates that only a small minority of organizations -- those McKinsey classifies as AI high performers, representing roughly six percent of respondents -- attribute a measurable share of enterprise earnings to AI deployment (McKinsey & Company, 2026b). This gap between infrastructure access and value realization is the organizing puzzle of this chapter.

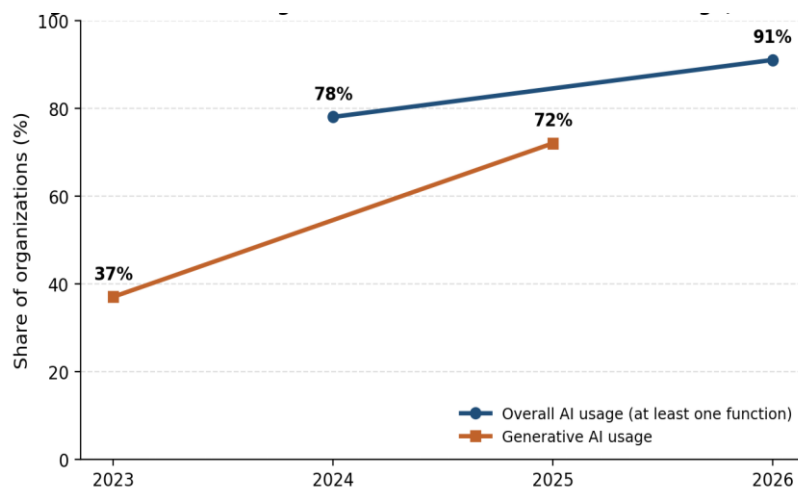


Figure 1: Growth in Organizational AI and Generative AI Usage, 2023-2026. Source: adapted from McKinsey & Company (2026a)

The remainder of the chapter proceeds as follows. Section 2 outlines the data and computing infrastructure underlying contemporary AI-driven smart systems. Section 3 introduces the theoretical lenses used to interpret adoption outcomes. Section 4 examines smart systems applications across core business functions. Section 5 addresses organizational enablers and barriers. Section 6 discusses governance, ethical, and security dimensions specific to autonomous and semi-autonomous smart systems. Section 7 situates the discussion within the labor-economics literature on AI and the future of work. Section 8 concludes with implications for practice and directions for future research.

2. Foundations: Data and Computing Infrastructure for Smart Business Systems

2.1 Defining AI-Enabled Smart Systems

A smart system, in the business management literature, refers to an integrated combination of sensors or data sources, computing infrastructure, and analytical or generative AI models that

together enable an organization to monitor conditions, generate predictions or recommendations, and, increasingly, take action with reduced human intervention (Collins, Dennehy, Conboy, & Mikalef, 2021). This definition distinguishes smart systems from earlier generations of business intelligence tools in two respects: smart systems typically operate on continuously updated, often real-time data rather than periodic batch reporting, and they increasingly incorporate a feedback loop in which system outputs directly influence operational action rather than only informing a subsequent human decision.

Four AI capability classes recur across the smart systems literature and structure the functional analysis in Section 4 of this chapter: automation, the execution of structured, repeatable tasks with minimal human involvement; prediction, the use of machine learning to forecast future states from historical and real-time data; personalization, the tailoring of content or recommendations to individual users at scale; and generative or agentic capability, the production of novel content or the execution of multi-step actions toward a defined goal with reduced step-by-step supervision. Each capability class depends on a different configuration of the underlying computing infrastructure discussed in the next section, summarized in Table 1.

2.2 The Computing Stack: Cloud, Edge, and Hybrid Architectures

Contemporary AI-driven smart systems typically draw on some combination of three computing architectures. Cloud computing provides centralized, elastically scalable processing power and remains the dominant infrastructure for training and hosting large-scale AI models, since cloud platforms have increasingly embedded AI and machine learning capability directly into core service offerings, with industry estimates suggesting a large majority of organizations now adopt AI at least partly through cloud platforms (Cyber Defense Magazine, 2025). Edge computing, by contrast, relocates processing closer to the point of data generation -- a factory floor sensor, a retail point-of-sale terminal, a connected vehicle -- reducing latency and bandwidth dependence for applications requiring near-instantaneous response, with industry forecasts projecting global spending on edge computing infrastructure to reach approximately 378 billion U.S. dollars by 2028 as enterprises prioritize localized compute capability (IDC, 2025).

Hybrid architectures, combining centralized cloud processing for computationally intensive model training with distributed edge processing for latency-sensitive inference, have become the dominant practical pattern for enterprise smart systems, particularly in manufacturing and logistics contexts where predictive maintenance and real-time quality control require millisecond-scale response unavailable through cloud-only architectures (Convex, 2025). Recent research on enterprise AI-cloud-edge integration frames this hybrid pattern as a holistic architecture for building sustainable, high-performance computing infrastructure capable of supporting the data-intensive applications and smart business systems driving contemporary digital transformation across sectors.

For business managers without deep technical backgrounds, the practical implication of this architecture discussion is a simple diagnostic: applications requiring real-time or near-real-time response -- fraud detection, predictive maintenance, autonomous logistics coordination -- generally require edge or hybrid infrastructure investment, while applications tolerant of periodic batch processing -- workforce planning, quarterly demand forecasting -- can typically rely on centralized cloud infrastructure alone, at meaningfully lower infrastructure cost and complexity.

2.3 Data Quality as the Binding Constraint

Regardless of computing architecture, the evidence synthesized throughout this chapter converges on a recurring finding: the quality, completeness, and consistency of the underlying data more often determines smart system performance than the sophistication of the AI model or the scale of the computing infrastructure applied to it. This finding follows directly from how contemporary machine learning systems function -- learning statistical patterns from historical examples -- since a system trained on incomplete or inconsistent data will faithfully reproduce those limitations in its outputs regardless of the underlying computing power available to it. Organizations planning smart systems investment should therefore budget explicitly for data governance and quality assurance alongside computing infrastructure, rather than treating data quality as a downstream implementation detail.

Table 1: AI Capability Classes and Their Computing Infrastructure Requirements

Capability Class	Primary Computing Requirement	Typical Latency Tolerance	Representative Application
Automation	Cloud or on-premises batch processing	Minutes to hours	Invoice processing, document classification
Prediction	Cloud-based model training; hybrid inference	Minutes to near real-time	Demand forecasting, credit risk scoring
Personalization	Cloud training; edge or hybrid inference	Seconds to sub-second	Real-time product recommendations
Generative / Agentic	High-throughput cloud (GPU-intensive)	Seconds	Content generation, autonomous workflow agents

3. Theoretical Perspectives on Smart Systems Adoption

Three theoretical lenses drawn from strategic management and information systems research anchor the analysis in this chapter. The resource-based view (RBV) of the firm holds that sustained competitive advantage arises from resources that are valuable, rare, inimitable, and non-substitutable (Barney, 1991). Applied to smart systems, the RBV implies that computing infrastructure and AI models, being commercially available to organizations of nearly any size,

rarely constitute a durable source of advantage on their own; advantage instead accrues to firm-specific complementary resources, particularly proprietary data assets and organizational capacity to redesign workflows around smart-systems outputs.

The Technology-Organization-Environment (TOE) framework organizes the determinants of technology adoption into technological factors (infrastructure, compatibility, complexity), organizational factors (leadership, skills, financial capacity, culture), and environmental factors (competitive pressure, regulation, institutional support) (Tornatzky & Fleischer, 1990). This framework is used in Section 5 to structure the discussion of smart systems adoption enablers and barriers, particularly the divide between large enterprises and small and medium-sized enterprises (SMEs).

Dynamic capabilities theory extends the RBV by emphasizing a firm's capacity to sense opportunities, seize them through resource reconfiguration, and continuously transform its resource base over time. This lens is particularly relevant to AI-driven smart systems because the underlying models require continuous retraining and the underlying infrastructure requires continuous reconfiguration as data volumes and computing demands evolve, making the capacity for organizational learning as important as the initial infrastructure investment.

4. Smart Systems across Core Business Management Functions

4.1 Strategic Decision Support

Smart systems increasingly support strategic decision-making through predictive and, increasingly, prescriptive analytics. A systematic review of peer-reviewed research on the convergence of AI and business strategy finds that value creation concentrates among firms that align AI investment explicitly with strategic objectives, while many firms mistake fragmented technological activity for coherent strategy, resulting in wasted resources and stalled initiatives. Emerging governance research on what has been termed the algorithmic boardroom examines how AI systems increasingly offer predictive and prescriptive strategic recommendations to boards and executive teams, with proposed governance frameworks emphasizing codified ethical values, board-level AI competency, and rigorous validation processes before AI-informed recommendations are acted upon.

4.2 Human Resource Management

Human resource management applications of AI-driven smart systems span recruitment, performance analytics, and workforce planning. Survey evidence from the Society for Human Resource Management indicates 43 percent of organizations now use AI in HR tasks, up from 26 percent the previous year (SHRM, 2025). Systematic reviews document AI's use in automating candidate screening and structuring interview analysis, while flagging significant unresolved concerns around fairness and legal compliance in automated candidate evaluation (Rigotti & Fosch-Villaronga, 2024). Case-study research on multinational corporations finds that AI-

enabled talent acquisition tools improve process efficiency but require substantial organizational investment in change management and bias auditing to be adopted successfully (Roppelt, Greimel, Kanbach, Stubner, & Maran, 2024).

4.3 Marketing and Customer-Facing Smart Systems

Marketing applications of smart systems center on real-time personalization and predictive consumer analytics. An integrative synthesis of empirically grounded studies on AI in customer service and strategic marketing identifies personalized interaction, automated service recovery, and dynamic communication as primary mechanisms of value creation, while highlighting a persistent tension between technological efficiency and perceived authenticity of engagement (Sahut & Laroche, 2025). This tension connects directly to the data infrastructure discussed in Section 2: real-time personalization at scale depends on low-latency data pipelines capable of incorporating customer behavior signals into recommendation systems within the timeframe of an active customer interaction, a requirement that has driven meaningful edge and hybrid computing investment within marketing technology stacks specifically.

4.4 Operations and Supply Chain Smart Systems

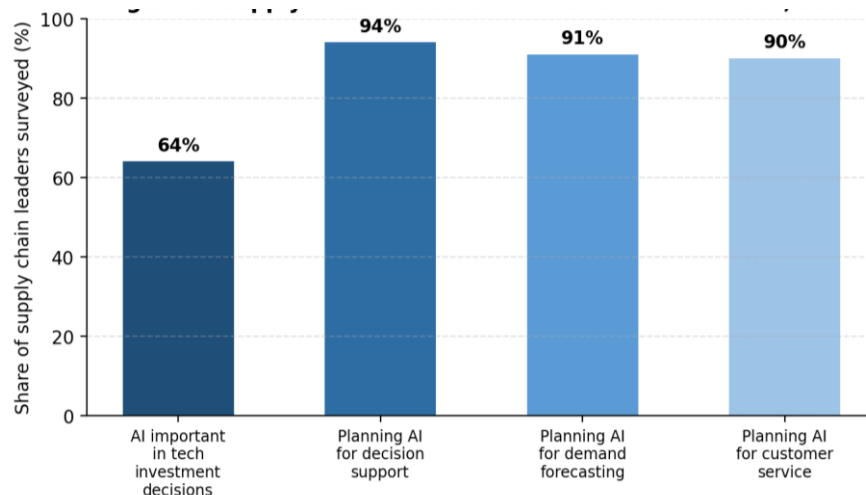


Figure 2: Supply Chain Leaders' AI Investment Priorities, 2025. Source: adapted from ABI Research (2025)

Operations and supply chain management represent among the most computing-infrastructure-intensive applications of smart systems in business, given the discipline's reliance on continuous sensor data from manufacturing equipment, logistics networks, and inventory systems. A systematic review examining AI integration across the transition from Industry 4.0 to Industry 6.0 finds substantial improvement in demand forecasting, inventory management, and supply chain decision-making capability, with an emerging emphasis on human-centric integration and sustainability alongside efficiency (Samuels, 2025). Industry survey data reinforce this pattern: a 2025 survey of 490 supply chain professionals found that 64 percent of supply chain leaders consider AI and generative AI capability important when evaluating new technology

investments, with 94 percent of organizations planning to use AI for decision support and over 90 percent planning to apply it to demand forecasting and customer service (ABI Research, 2025), as illustrated in Figure 2.

These operations applications depend disproportionately on the edge and hybrid computing architectures discussed in Section 2.2, since predictive maintenance and real-time quality control require the low-latency processing edge infrastructure specifically provides, distinguishing this functional domain from the more cloud-centric infrastructure requirements typical of the marketing and strategic applications discussed above.

4.5 Finance, Accounting, and Risk Management

In finance, AI-driven smart systems concentrate on fraud detection, credit-risk assessment, and algorithmic decision support for capital allocation. Machine-learning-based approaches to financial risk assessment demonstrate improved early identification of default risk relative to traditional financial-ratio-based models, particularly under conditions of macroeconomic uncertainty. Firm-level research using U.S. data finds AI adoption associated with measurable increases in firm growth and product innovation output, with effects concentrated among firms combining AI adoption with complementary investment in skilled labor and organizational restructuring (Babina, Fedyk, He, & Hodson, 2024) -- a finding directly consistent with the resource-based view argument introduced in Section 3.

5. Organizational Enablers and Barriers to Smart Systems Adoption

Despite near-universal experimentation with AI-driven smart systems, evidence consistently points to a substantial gap between infrastructure adoption and realized business value. This section applies the TOE framework introduced in Section 3 to structure the discussion of adoption determinants.

5.1 Technological and Organizational Readiness

Technological readiness -- computing infrastructure, data pipeline maturity, and system integration complexity -- remains a frequently cited barrier, particularly outside large enterprises. However, organizational factors emerge across the literature as more decisive than technological readiness alone in determining outcomes. McKinsey's 2026 State of Organizations survey of over 10,000 respondents identifies ease of use, leadership sponsorship, and dedicated implementation teams as key enablers of enterprise-wide AI adoption, illustrated by the case of the German insurer Allianz, whose internally developed generative AI platform scaled to more than 60,000 active users across underwriting, claims, and product design functions (McKinsey & Company, 2026c). Organizations McKinsey classifies as AI high performers are nearly three times as likely to have fundamentally redesigned workflows around AI and more than twice as likely to invest over fifteen percent of their enterprise-wide technology budget in AI, compared with other respondents (McKinsey & Company, 2026b).

Table 2: TOE Framework Dimensions in Smart Systems Adoption, by Firm Size

TOE Dimension	Large Enterprise Pattern	SME Pattern
Technological	Mature cloud/edge infrastructure; dedicated AI engineering teams	Legacy systems; limited data infrastructure; low familiarity with generative AI
Organizational	Dedicated budgets; structured change management; executive sponsorship	Resource constraints; limited in-house technical expertise
Environmental	Greater capacity to absorb regulatory compliance costs	High sensitivity to competitor moves; reliance on external support programs

5.2 The SME Computing Infrastructure Gap

Research applying an integrated TOE-resource-based-view lens to SMEs operating in resource-constrained environments finds that technological readiness and leadership vision function as strategic resources mediating the relationship between AI adoption and firm performance, suggesting that policy interventions strengthening these specific capacities are likely more effective in narrowing the adoption gap than subsidies for AI tool acquisition alone. The rise of cloud-based, consumption-priced computing infrastructure has partially lowered the capital barrier historically associated with enterprise-grade computing, but SMEs continue to report limited familiarity with configuring and integrating this infrastructure effectively, indicating the binding constraint has shifted from capital access toward technical literacy.

6. Governance, Ethics, and Security in Autonomous Smart Systems

As smart systems increasingly incorporate autonomous or semi-autonomous decision agents, governance considerations extend beyond the data-privacy and algorithmic-bias concerns associated with earlier AI applications to encompass system-level questions about autonomous action in physical and operational environments. Research on AI-powered business analytics finds that without deliberate fairness-auditing processes, AI models trained on historical organizational data can systematically reproduce and amplify existing patterns of discrimination in hiring, lending, and customer treatment. Recent frameworks proposed for board-level AI governance emphasize three complementary pillars: explicit codification of organizational ethical principles within AI systems; investment in the human competency required for boards and executives to meaningfully oversee AI-informed decisions; and rigorous technical validation processes prior to deployment of AI systems in consequential decision domains.

Security considerations specific to distributed, edge-heavy smart systems architectures deserve particular attention: the proliferation of edge computing nodes, each representing a potential point of vulnerability, expands an organization's attack surface relative to centralized cloud-only architectures, requiring dedicated security governance -- device authentication, encrypted data

transmission, and continuous monitoring across distributed nodes -- that general-purpose AI governance frameworks do not always address explicitly. Organizations deploying edge-heavy smart systems in operations and supply chain contexts, discussed in Section 4.4, should treat this distributed security governance as a distinct requirement layered atop the general AI governance frameworks discussed throughout this section.

7. Economic and Workforce Implications

The organizational transformations examined throughout this chapter intersect with broader labor-market dynamics. A foundational strand of labor-economics research, developed principally by Daron Acemoglu and David Autor, models AI's labor-market effects through a task-based framework distinguishing between the displacement effect, whereby AI automates tasks previously performed by workers, and the reinstatement effect, whereby new technologies create new categories of task (Acemoglu & Restrepo, 2022). Using data on U.S. online vacancy postings, this research finds measurable evidence of both effects operating simultaneously in occupations exposed to AI (Acemoglu, Autor, Hazell, & Restrepo, 2022). More recent work extends this framework specifically to generative AI, arguing that AI's most consequential labor-market effect may be its capacity to extend the reach of human expertise rather than to automate it away entirely (Autor, 2024).

This literature intersects with a broader interdisciplinary body of work examining AI's societal and economic implications. Research on the socioeconomic impact of AI across labor markets and institutional structures characterizes AI's economic and social consequences as substantially dependent on the governance and deployment choices made by adopting organizations, a conclusion consistent with the organizational-primacy argument developed throughout this chapter (Mishra, Mishra, & Agarwal, 2026a). Related work on the concentration of AI and computing infrastructure across firms and geographies carries direct relevance to the computing-infrastructure discussion in Section 2, arguing that unequal access to advanced computing resources -- particularly the GPU-intensive infrastructure required for generative and agentic AI, discussed in Table 1 -- has implications for competitive dynamics between organizations of different scale that parallel the SME-versus-large-enterprise divide documented in Section 5 (Mishra, Mishra, & Agarwal, 2026b).

8. Illustrative Applications in Smart Systems Deployment

8.1 Smart Manufacturing

The manufacturing sector offers a useful illustration of the integrated argument developed throughout this chapter, since it combines all three computing architectures, all four AI capability classes, and the governance and workforce considerations discussed in Sections 6 and 7 within a single operational environment. A composite scenario, consistent with the digital-twin and predictive-maintenance applications discussed in the operations literature cited in Section

4.4: a mid-sized manufacturer deploys edge-based sensors across production equipment, feeding a hybrid cloud-edge architecture that performs real-time anomaly detection at the edge while sending aggregated data to the cloud for longer-horizon predictive-maintenance model retraining. The initiative succeeds not because the underlying sensors or models are unusually sophisticated, but because the manufacturer pairs the infrastructure investment with a redesigned maintenance workflow, retrained maintenance staff, and a governance process for validating model-generated maintenance alerts before equipment shutdown decisions are made -- illustrating, in a single operational example, this chapter's recurring argument that infrastructure and organizational redesign are complementary, not substitutable, investments.

8.2 Smart Retail and Real-Time Personalization

A second illustration, drawn from the marketing and customer-facing applications discussed in Section 4.3: a retail organization deploys edge computing within physical store locations to support real-time inventory visibility and personalized in-store promotions delivered to customers' mobile devices, while relying on centralized cloud infrastructure for the less latency-sensitive task of longer-horizon demand forecasting across its full store network. This architecture choice reflects the latency-based diagnostic introduced in Section 2.2: the in-store personalization application, requiring sub-second response to a customer's real-time location and browsing behavior, justifies edge investment that the demand-forecasting application, tolerant of daily or weekly batch updates, does not require, allowing the retailer to allocate computing infrastructure investment efficiently across applications with genuinely different latency requirements rather than defaulting to a uniform infrastructure approach across all use cases.

Conclusion and Future Research Directions

This chapter has examined AI-driven smart systems in business management through an integrated lens spanning computing infrastructure, data governance, organizational adoption theory, functional application, and workforce implications. Three conclusions merit emphasis. First, computing infrastructure and data quality are necessary but not sufficient conditions for smart-systems value creation; organizational redesign and governance maturity remain the decisive determinants of whether infrastructure investment converts into measurable business value, consistent with evidence that only a small minority of organizations achieve high-performer status despite near-universal infrastructure adoption. Second, computing architecture choices -- cloud, edge, or hybrid -- are not merely technical decisions but carry direct implications for which business functions and applications an organization can support effectively, with latency-sensitive operations and marketing applications depending disproportionately on edge and hybrid infrastructure investment. Third, the organizational and workforce dimensions of smart-systems adoption are inseparable from the technical dimensions emphasized elsewhere in this volume: how organizations deploy AI-driven smart systems shapes

not only operational efficiency but also workforce outcomes and competitive dynamics between organizations with unequal access to advanced computing resources.

Future research would benefit from longitudinal study of the specific mechanisms linking computing architecture choices to functional business outcomes, deeper comparative research on SME computing infrastructure adoption across regional and institutional contexts, and continued interdisciplinary work connecting the technical computing and data science literature represented elsewhere in this volume with the organizational and labor-economics literature emphasized in this chapter. As computing and data infrastructure continue to commoditize, per the discussion in Section 2, the organizational and governance questions addressed in Sections 5 through 7 are likely to become, if anything, more rather than less central to how the value of smart systems in business management is ultimately determined and studied.

Acknowledgements

The authors declare no conflicts of interest. We are very much thankful to the authors of different publications as many new ideas are abstracted from them. Authors also express gratefulness to their colleagues and family members for their continuous help, inspirations, encouragement, and sacrifices without which this work could not be executed. Finally, the main target of this work will not be achieved unless it is used by research institutions, students, research scholars, and authors in their future works. The authors will remain ever grateful to Dr. Neelu Singh, Director, ICFRE Tropical Forest Research Institute, Jabalpur, Director, XLRI – Xavier School of Management, Jamshedpur & Principal Government Science College, Jabalpur who helped by giving constructive suggestions for this work. The authors are also responsible for any possible errors and shortcomings, if any in the paper, despite the best attempt to make it immaculate.

References

1. ABI Research. (2025). *2025 supply chain survey results: Artificial intelligence (AI) usage and investment plans*. ABI Research.
2. Acemoglu, D., & Restrepo, P. (2022). Tasks, automation, and the rise in U.S. wage inequality. *Econometrica*, 90(5), 1973–2016.
3. Acemoglu, D., Autor, D., Hazell, J., & Restrepo, P. (2022). Artificial intelligence and jobs: Evidence from online vacancies. *Journal of Labor Economics*, 40(S1), S293–S340.
4. Autor, D. (2024). Applying AI to rebuild middle class jobs (NBER Working Paper No. 32140). National Bureau of Economic Research.
5. Babina, T., Fedyk, A., He, A., & Hodson, J. (2024). Artificial intelligence, firm growth, and product innovation. *Journal of Financial Economics*, 151, 103745.
6. Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99–120.

7. Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, 60, 102383.
8. Convox. (2025). *The AI-edge computing boom: Why infrastructure matters in 2025*. Convox Engineering Blog.
9. Cyber Defense Magazine. (2025). *The impact of cloud computing in 2025*. Cyber Defense Magazine.
10. IDC. (2025). *Worldwide edge enterprise infrastructure forecast, 2025–2028*. International Data Corporation.
11. McKinsey & Company. (2026a). *The state of AI: Global survey 2026*. QuantumBlack, AI by McKinsey.
12. McKinsey & Company. (2026b). *The state of AI: Global survey 2026 (enterprise ROI and workforce findings)*. QuantumBlack, AI by McKinsey.
13. McKinsey & Company. (2026c). *The state of organizations 2026*. McKinsey & Company.
14. Mishra, D., Mishra, R. K., & Agarwal, R. (2026). The socioeconomic impact of artificial intelligence on society. *International Journal of Recent Development in Engineering and Technology*, 15(3), 1477–1493.
15. Mishra, D., Mishra, R. K., & Agarwal, R. (2026). Power concentration and the political economy of AI. In *Co-Intelligence: Human-AI Coexistence in the Age of Thinking Machines* (pp. 96–99). Deep Science Publishing. E-Book ISBN-13: 978-93-7185-166-4; Paperback ISBN-13: 978-93-7185-178-7. <https://doi.org/10.70593/978-93-7185-166-4>
16. Rigotti, C., & Fosch-Villaronga, E. (2024). Fairness, AI & recruitment. *Computer Law & Security Review*, 53, 105966.
17. Roppelt, J. S., Greimel, N. S., Kanbach, D. K., Stubner, S., & Maran, T. K. (2024). Artificial intelligence in talent acquisition: A multiple case study on multinational corporations. *Management Decision*, 62(10), 2986–3007.
18. Sahut, J. M., & Laroche, M. (2025). Using artificial intelligence (AI) to enhance customer experience and to develop strategic marketing: An integrative synthesis. *Computers in Human Behavior*, 170, 108684.
19. Samuels, A. (2025). Examining the integration of artificial intelligence in supply chain management from Industry 4.0 to 6.0: A systematic literature review. *Frontiers in Artificial Intelligence*. <https://doi.org/10.3389/frai.2024.1477044>
20. Society for Human Resource Management [SHRM]. (2025). *2025 talent trends: Artificial intelligence in HR*. SHRM Research.
21. Tornatzky, L. G., & Fleischer, M. (1990). *The processes of technological innovation*. Lexington Books.

STATISTICAL METHODS FOR COMPUTING AND DATA SCIENCE: FOUNDATIONS, ALGORITHMS, AND APPLICATIONS

Umesh R. Yeole

Department of Statistics,

Pratap College, (Autonomous) Amalner, 425401, Maharashtra, India

Corresponding author E-mail: uryeole@gmail.com

Abstract

The rapid growth of computing power and the explosion of data across every domain have made statistics not merely a supporting discipline but a cornerstone of modern data science. This chapter presents a structured, end-to-end treatment of the statistical methods that underpin computational data analysis, organized around three interconnected pillars: foundations, algorithms, and applications. The foundations section establishes the theoretical bed rock probability theory, statistical inference, Bayesian reasoning, and information theory while highlighting the computational considerations that distinguish modern statistical practice from classical theory. Building on this base, the chapter surveys core statistical methods, including regression and generalized linear models, dimensionality reduction, classification and clustering, regularization, resampling, nonparametric techniques, and time series analysis. It then examines how these methods are realized as algorithms, covering optimization strategies, Monte Carlo and Markov chain Monte Carlo techniques, variational inference, numerical linear algebra, and approaches for scaling inference to large and high-dimensional datasets. The applications section demonstrates how these methods translate into practice through case studies in healthcare, finance, natural language processing, computer vision, recommender systems, and social network analysis. Finally, the chapter addresses contemporary challenges including fairness and ethics, uncertainty quantification, causality, and privacy-preserving statistics—and outlines emerging directions such as Bayesian deep learning, conformal prediction, and federated learning. By integrating theory, computation, and real-world application, the chapter aims to equip students, researchers, and practitioners with both the conceptual understanding and the practical toolkit needed to apply statistical methods rigorously in computing and data science.

Introduction

The last two decades have witnessed a fundamental transformation in how knowledge is extracted from data. Advances in computing power, the proliferation of digital sensors, and the near-ubiquity of the internet have produced datasets of a scale and complexity that would have been unimaginable to earlier generations of statisticians. Yet the methods used to draw reliable conclusions from these data remain, at their core, statistical. Machine learning models are trained

by optimizing likelihoods and minimizing expected loss; recommendation systems rely on probabilistic inference over user behavior; clinical trials, financial risk models, and genomic studies all depend on rigorous statistical reasoning. In short, computing has provided the *means* to process data at scale, but statistics provides the *logic* by which those data yield trustworthy knowledge (Botev *et al.*, 2019; Fan *et al.*, 2020).

This interdependence is precisely why a chapter devoted to statistical methods in computing and data science is both timely and necessary. Too often, the statistical foundations of modern data analysis are treated as background material assumed rather than taught, glossed over in the rush to implement the latest algorithm. The consequence is a generation of practitioners who can fit models but cannot always explain *why* they work, *when* they fail, or *how much* confidence to place in their predictions (Henley *et al.*, 2020; Wang *et al.*, 2016). This chapter seeks to close that gap by presenting statistical methods as an integrated whole: theory that motivates practice, algorithms that operationalize theory, and applications that demonstrate their value.

Statistical methods for computing and data science rest on three linked pillars: inferential foundations, scalable algorithms, and domain applications, with modern practice increasingly shaped by high dimensionality, large sample size, and computational constraints (Abdulah *et al.*, 2025; Jordan, 2013; Volodina & Challenor, 2021a). The literature also converges on a central theme: useful modern methods must balance statistical accuracy with computational efficiency, reproducibility, and uncertainty quantification rather than treating these as separate concerns (Abdulah *et al.*, 2025; Sohil *et al.*, 2022; Volodina & Challenor, 2021b).

Statistical methods in computing and data science aim to turn data into prediction, inference, and decisions by combining probabilistic reasoning, statistical modeling, and computational algorithms. This synthesis has become central because modern scientific and industrial systems generate data at scales that force statistics and computer science into closer interaction.

The modern field spans supervised learning, unsupervised learning, regression, classification, clustering, dimensionality reduction, kernel methods, ensemble methods, support vector machines, neural networks, and deep learning (Botev *et al.*, 2019; Bouveyron *et al.*, 2019). Across these areas, the practical goal is not only to fit models, but to understand when methods generalize, how they scale, what assumptions they require, and how their outputs should be interpreted under uncertainty (Rocklöv & Gayle, 2021; Wu, 1999).

The remainder of the chapter proceeds as follows. Section 2 lays the theoretical foundations probability, statistical inference, Bayesian reasoning, information theory, and key computational considerations. Section 3 surveys core statistical methods, from regression and dimensionality reduction to clustering, regularization, and time series analysis. Section 4 examines the algorithms that implement these methods, including optimization, Monte Carlo techniques, variational inference, and scaling strategies for big data. Section 5 presents applications across

several domains, illustrating how theory and computation combine in practice. Section 6 discusses contemporary challenges fairness, uncertainty quantification, causality, and privacy and outlines emerging directions. Section 7 concludes with a summary and practical guidance. Together, these sections give readers both a conceptual map and a working toolkit for applying statistical methods in computing and data science.

Foundation

This section establishes the theoretical bedrock on which all subsequent methods rest. It covers probability essentials, statistical inference, Bayesian reasoning, information theory, and the computational considerations that distinguish modern statistical practice from its classical origins. Readers already comfortable with these topics may skim, but the material here provides the vocabulary and intuition used throughout the chapter.

Probability Essentials

A random variable X maps outcomes of an experiment to numerical values, and its behavior is described by a probability distribution. For discrete variables, the probability mass function $p(x) = P(X = x)$ assigns probabilities to individual values; for continuous variables, the probability density function $p(x)$ describes relative likelihood over intervals, with probabilities obtained by integration.

Two summaries dominate practice. The expectation $E[X]$ captures the long-run average, while the variance $Var(X)$ measures dispersion around that average. These extend naturally to multiple variables through joint distributions $p(x, y)$, marginal distributions $p(x) = \int p(x, y) dy$, and conditional distributions $p(y | x) = p(x, y)/p(x)$. Independence where $p(x, y) = p(x)p(y)$ is a special and consequential case (Ibe, 2014; Jacod & Protter, 2004).

Two limit theorems anchor statistical practice. The law of large numbers guarantees that sample averages converge to their theoretical expectations as sample size grows, justifying empirical estimation. The central limit theorem states that sums (and averages) of independent random variables tend toward a normal distribution, regardless of the underlying distribution a result that underpins confidence intervals, hypothesis tests, and countless modeling assumptions.

Common distributions recur throughout the chapter: the Bernoulli and binomial for binary outcomes, the Poisson for counts, the uniform for bounded intervals, the normal (Gaussian) for continuous measurements, and the exponential for waiting times. Familiarity with these families and their properties is assumed, though key results are recalled where used (Rényi, 2007).

Statistical Inference Fundamentals

Inference is the process of drawing conclusions about unknown population parameters from observed samples. Formally, we posit a statistical model $p(x | \theta)$ indexed by a parameter θ , and seek to estimate θ , test hypotheses about it, or quantify uncertainty in our conclusions.

An estimator $\hat{\theta}$ is a function of the data intended to approximate θ . Its quality is assessed through several properties: bias (systematic deviation from the true value), variance (spread across samples), mean squared error ($\text{bias}^2 + \text{variance}$), consistency (convergence to the truth as data grows), and efficiency (achieving low variance relative to theoretical bounds such as the Cramér–Rao lower bound).

The most widely used estimation principle is maximum likelihood estimation (MLE). Given independent observations $x_1, \dots, x_n$, the likelihood is $L(\theta) = \prod_{i=1}^n p(x_i|\theta)$ and the MLE is $\hat{\theta} = \operatorname{argmax}_{\theta} \operatorname{Log}(L(\theta))$. Under regularity conditions, MLEs are consistent, asymptotically normal, and asymptotically efficient properties that explain their ubiquity.

Hypothesis testing formalizes decisions under uncertainty. A null hypothesis H_0 is compared against an alternative H_1 using a test statistic whose distribution under H_0 is known. The p-value is the probability, assuming H_0 , of observing a statistic at least as extreme as that obtained; small p-values suggest evidence against H_0 . The significance level α controls the Type I error rate (false positives), while power $1 - \beta_1$ reflects the ability to detect true effects (avoiding false negatives). Confidence intervals provide a complementary view, giving a range of parameter values consistent with the data at a specified coverage level (Casella & Berger, 2024).

These classical tools remain indispensable, but they carry assumptions independence, correct model specification, adequate sample size that must be checked in practice.

Bayesian Foundations

The Bayesian paradigm treats parameters as random variables with prior distributions $p(\theta)$ reflecting beliefs before observing data. After observing data x , beliefs are updated via Bayes' theorem

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}$$

Where $p(\theta|x)$ is the posterior, $p(x|\theta)$ is the likelihood, and $p(x) = \int p(x|\theta)p(\theta)d\theta$ represents the likelihood or marginal likelihood. The posterior combines prior knowledge with data and serves as the basis for all subsequent inference point estimates (e.g., the posterior mean or mode), credible intervals, and predictive distributions.

The Bayesian approach offers several advantages: it quantifies uncertainty directly through probability statements about parameters, naturally incorporates prior information, and extends gracefully to hierarchical and complex models. Its principal challenge is computational: the evidence integral is often intractable, motivating the Monte Carlo and variational methods discussed in Section 4 (Gamerman *et al.*, 2026).

The frequentist–Bayesian distinction is worth noting explicitly. Frequentists treat parameters as fixed unknowns and probabilities as long-run frequencies; Bayesians treat parameters as random and probabilities as degrees of belief. Both frameworks are used throughout the chapter, and

readers are encouraged to appreciate their complementary strengths rather than treat them as competitors.

Information Theory Basics

Information theory provides a bridge between statistics and machine learning. The entropy of a random variable

$$H(X) = - \sum_x p(x) \log p(x)$$

quantifies its average uncertainty. The Kullback–Leibler (KL) divergence,

$$D_{KL}(p||q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$$

measures how much one distribution q diverges from another p ; it is non-negative and zero only when $p = q$. Mutual information $I(X; Y) = H(X) - H(X|Y)$ captures the dependence between two variables (Cover *et al.*, 1991; Gray, 2011).

These quantities appear repeatedly in later sections: cross-entropy loss in classification is a KL divergence in disguise; maximum likelihood estimation is equivalent to minimizing KL divergence from the empirical distribution; and variational inference (Section 4) explicitly minimizes a KL divergence between an approximate and true posterior (Ash, 1990).

Computational Considerations

Classical statistical theory often assumes exact arithmetic and moderate dimensions. Modern practice violates both assumptions, and a working knowledge of computational pitfalls is essential.

Numerical stability is a first concern. Probabilities can underflow when multiplied across many observations, which is why likelihoods are computed in log space. Matrix inversions can amplify rounding errors, motivating decompositions (QR, Cholesky, SVD) that are more stable. Summations of many small terms may lose precision, and catastrophic cancellation can silently corrupt results.

The curse of dimensionality is a second concern. As the number of features d grows, the volume of the space explodes, distances between points become less informative, and the sample size required for reliable estimation grows exponentially. Methods such as regularization, dimensionality reduction, and sparsity-inducing priors (Sections 3 and 4) exist precisely to combat this.

Computational complexity is a third concern. An algorithm that is theoretically optimal may be impractical if it scales as $O(n^3)$ in the sample size n . Conversely, approximate methods with favorable scaling may deliver sufficient accuracy at a fraction of the cost. Throughout the chapter, complexity is discussed alongside statistical properties, since the two are inseparable in practice (Gabriel & Marco, 2019).

Finally, reproducibility deserves mention even at the foundational level. Random seeds, numerical tolerances, and software versions all affect results; a statistical claim that cannot be reproduced is of limited scientific value. Section 4 revisits this theme in the context of software engineering.

Core Statistical Methods

This section surveys the principal statistical methods in computing and data science, emphasizing what each does, its assumptions, and when to use it. Implementation details appear in Section 4; applications in Section 5.

Regression and Generalized Linear Model

Regression models the relationship between a response Y and the predictors $X_1, X_2, \dots, X_p$. Linear regression assumes $Y = X\beta + \epsilon$ and is fit by ordinary least squares (OLS). Generalized linear models (GLMs) extend this to non-normal responses via a link function $g(\cdot): E[Y|X] = g^{-1}(X\beta)$ (Sykes, 1993).

Table 1: Common Regression Models

Model	Response	Link	Use
Linear	Continuous	Identity	Prediction, inference
Logistic	Binary	Logit	Classification, odds ratios
Poisson	Count	Log	Event rates
Multinomial	Categorical	Softmax	Multi-class

Dimensionality Reduction

High-dimensional data often lie near a lower-dimensional manifold. PCA finds orthogonal linear combinations maximizing variance. Factor analysis models latent factors with noise. t-SNE and UMAP produce nonlinear embeddings for visualization (Fodor, 2002).

Table 2: Dimensionality Reduction Methods

Method	Linearity	Preserves	Best For
PCA	Linear	Global variance	Compression
Factor analysis	Linear	Latent structure	Latent traits
t-SNE	Nonlinear	Local neighborhoods	Visualization
UMAP	Nonlinear	Local + global	Embeddings
Method	Linearity	Preserves	Best For

Classification and Clustering

Classification assigns labels using methods like logistic regression, LDA/QDA, naive Bayes, and k-NN. Clustering discovers groups via k-means, Gaussian mixture models (GMMs), hierarchical clustering, and DBSCAN (Arabie *et al.*, 1996; Lavine & Mirjankar, 2000; Mirkin, 2013).

Table 3: Classification and Clustering Method

Method	Task	Assumption
Logistic regression	Classification	Linear log-odds
LDA / QDA	Classification	Gaussian classes
Naive Bayes	Classification	Feature independence
k-NN	Classification	Local smoothness
k-means	Clustering	Spherical clusters
GMM	Clustering	Gaussian mixtures
DBSCAN	Clustering	Density-based

Resampling and Model Validation

Cross-validation partitions data into k folds to estimate test error. The bootstrap resamples with replacement for confidence intervals and standard errors. Stratified and time-series CV handle imbalanced and temporal data respectively (James *et al.*, 2023).

Table 4: Resampling Methods

Method	Purpose	Notes
k-fold CV	Estimate test error	Standard choice
LOOCV	Unbiased estimate	High variance
Bootstrap	Sampling distribution	CI, standard errors
Jackknife	Bias/variance	Leave-one-out

Regularization and Variable Selection

Regularization trades bias for variance when predictors are many or correlated. Ridge ℓ_2 shrinks coefficients; LASSO ℓ_1 selects variables; elastic net combines both (Zou & Hastie, 2005).

Table 5: Regularization Methods

Method	Penalty	Effect
Ridge	ℓ_2	Shrinks coefficients
LASSO	ℓ_1	Sparse selection
Elastic net	$\ell_1 + \ell_2$	Sparse + stable

Nonparametric Methods

Nonparametric methods relax functional assumptions: kernel density estimation, splines/GAMs, rank-based tests, and LOESS. They are flexible and robust but data-hungry (Sijtsma & Emons, 2010).

Time Series and Sequential Methods

ARIMA captures trend and autocorrelation; SARIMA adds seasonality. State-space models and the Kalman filter handle latent dynamics. Exponential smoothing offers simple forecasting (Brockwell & Davis, 2009).

Table 6: Time Series Methods

Method	Captures
ARIMA	Trend, autocorrelation
SARIMA	Seasonality + ARIMA
Kalman filter	Latent state
Exponential smoothing	Level, trend, season

Algorithms and Computation

Section 3 presented statistical methods conceptually. This section examines how they are implemented, optimized, and scaled turning theory into working code. Topics include the statistical computing landscape, optimization, Monte Carlo methods, numerical linear algebra, scaling to big data, and reproducibility.

The Statistical Computing Landscape

Modern statistical work is done in a small number of languages and ecosystems. Python dominates industry through NumPy (arrays), SciPy (scientific computing), pandas (data frames), statsmodels (classical statistics), and scikit-learn (machine learning). R remains strong in statistics and biostatistics. The choice often depends on the surrounding pipeline rather than the statistics alone (Abdulah *et al.*, 2025).

Table 7: Statistical Computing Environments

Environment	Strengths
R	Rich statistical libraries, visualization
Python	General-purpose, ML ecosystem
Julia	Speed, numerical computing
MATLAB	Numerical linear algebra
SQL	Data manipulation at scale

Optimization for Statistics

Almost every statistical method reduces to an optimization problem: maximizing likelihood, minimizing loss, or finding a posterior mode.

Gradient descent iteratively updates parameters in the direction of the negative gradient. Stochastic gradient descent (SGD) uses mini-batches, enabling large-scale learning. Newton–Raphson uses second-order information for faster convergence near the optimum. The Expectation–Maximization (EM) algorithm alternates between estimating latent variables (E-step) and updating parameters (M-step), and underlies GMMs, factor analysis, and hidden Markov models. Coordinate descent optimizes one parameter at a time and is particularly effective for LASSO-type problems (Everitt, 2012).

Table 8: Optimization Methods

Method	Order	Best For
Gradient descent	First	Smooth convex problems
SGD / mini-batch	First	Large-scale learning
Newton–Raphson	Second	Small–medium problems
EM	—	Latent-variable models
Coordinate descent	First	LASSO, sparse models
L-BFGS	Quasi-Newton	General smooth problems

Monte Carlo Methods

Many statistical quantities posterior expectations, marginal likelihoods, complex integrals cannot be computed analytically. Monte Carlo methods approximate them by simulation.

Basic Monte Carlo estimates $E[f(X)]$ by averaging $f(x_i)$ over samples $x_i \sim p(x)$. Importance sampling draws from a proposal distribution and reweights, useful when p is hard to sample directly. Markov chain Monte Carlo (MCMC) constructs a Markov chain whose stationary distribution is the target; Metropolis Hastings and Gibbs sampling are the classic algorithms. Hamiltonian Monte Carlo (HMC) and its variant NUTS use gradient information for efficient exploration and underlie modern probabilistic programming tools (Gamerman *et al.*, 2026; Hammersley, 2013; Kroese & Rubinstein, 2012).

Variational inference (VI) offers a faster alternative: instead of sampling, it approximates the posterior with a tractable family $q(\theta)$ and minimizes the KL divergence to the true posterior equivalently, maximizes the evidence lower bound (ELBO). VI is faster than MCMC but typically less accurate.

Table 9: Monte Carlo and Approximate Inference Methods

Method	Idea	Speed	Accuracy
Monte Carlo	Sample and average	Fast	Moderate
Importance sampling	Reweighted samples	Fast	Depends on proposal
Metropolis–Hastings	Random-walk MCMC	Slow	High
Gibbs sampling	Conditional sampling	Moderate	High
HMC / NUTS	Gradient-based MCMC	Moderate	High
Variational inference	Optimize approximate posterior	Fast	Approximate

Applications

Sections 3 and 4 presented methods and algorithms in the abstract. This section demonstrates how they combine in practice across six domains. Each subsection follows a consistent pattern: the problem, the statistical methods used, the computational approach, and the lessons learned.

Healthcare and Bioinformatics

Clinical decision-making, drug discovery, and genomic analysis all require reliable inference from noisy, high-dimensional, and often sensitive data. Statistical methods are central to this

enterprise. Logistic regression and survival analysis, particularly Cox proportional hazards models, are standard tools for modeling clinical outcomes and time-to-event data, while mixed-effects models accommodate the longitudinal structure of patient records collected over repeated visits. In genomics, where tens of thousands of gene expression measurements are available per sample, principal component analysis and clustering are used to uncover latent structure and identify disease subtypes. Bayesian hierarchical models are especially valuable in multi-site clinical trials, where they pool information across sites while respecting site-level variation. Computationally, the high dimensionality of genomic data demands regularization techniques such as LASSO and elastic net to select relevant features and avoid overfitting, and growing privacy concerns have motivated federated learning approaches that allow hospitals to collaborate without sharing raw patient data. The overarching lesson from this domain is that interpretability and calibrated uncertainty matter as much as predictive accuracy; clinicians and regulators must be able to understand and trust a model's outputs, and this trust depends on reproducible, auditable analytical pipelines (Mishra *et al.*, 2024; Verma *et al.*, 2024).

Finance and Risk Modeling

Financial institutions must quantify risk, detect fraud, and price assets under uncertainty, making statistics indispensable to the industry. Time series models such as ARIMA and GARCH are used to model asset prices and volatility, with GARCH particularly effective at capturing the volatility clustering observed in financial markets. Logistic regression and gradient boosting are widely deployed for credit scoring and fraud detection, where the goal is to separate legitimate from suspicious activity in highly imbalanced datasets. For tail risk, where extreme events dominate losses, copulas and extreme value theory provide tools for modeling dependence and rare-event behavior that Gaussian assumptions cannot capture. Monte Carlo simulation remains the workhorse for pricing complex, path-dependent derivatives. Computationally, real-time fraud detection demands low-latency inference on streaming data, while risk simulations rely on large-scale parallel Monte Carlo. The key lesson is that heavy-tailed distributions and non-stationarity routinely break naive Gaussian assumptions, and backtesting and stress testing are essential before any model is deployed in production (Aldridge & Avellaneda, 2021; Cao *et al.*, 2021; Tatsat *et al.*, 2020).

Natural Language Processing

Text data is discrete, sparse, high-dimensional, and deeply context-dependent, presenting distinctive statistical challenges. Naive Bayes and logistic regression have long served as interpretable baselines for text classification tasks such as spam detection and sentiment analysis. Latent Dirichlet *et al.* location provides a probabilistic framework for discovering topics in document collections, while word embeddings such as word2vec and GloVe learn dense vector representations that capture semantic similarity. Sequence models, including hidden Markov models and conditional random fields, are used for structured prediction tasks like part-of-speech

tagging and named entity recognition. Computationally, sparse matrix representations, stochastic optimization, and distributed training enable these methods to scale to billions of tokens. The lesson here is that statistical foundations remain relevant even in the era of large language models; probabilistic modeling informs uncertainty, calibration, and evaluation in ways that purely predictive systems often overlook (Lane & Dyshe, 2025; Zeroual & Lakhouaja, 2018).

Computer Vision and Signal Processing

Images and signals are high-dimensional, spatially correlated, and frequently corrupted by noise. Classical statistical methods continue to underpin much of the preprocessing and low-level analysis in this domain. Gaussian filtering and the Kalman filter are used for denoising and object tracking, with the Kalman filter providing optimal linear state estimation under Gaussian assumptions. Principal component analysis, through the well-known eigenfaces approach, was an early milestone in face recognition, and Gaussian mixture models remain useful for background subtraction in video analysis. Graphical models such as Markov random fields and conditional random fields provide a principled framework for structured prediction, capturing spatial dependencies between neighboring pixels. Computationally, convolutions, fast Fourier transforms, and GPU acceleration make large-scale image processing feasible. The key lesson is that classical statistical methods remain the backbone of preprocessing and low-level vision, even as deep learning has come to dominate high-level tasks such as object detection and image segmentation (Little, 2019; Ricaud *et al.*, 2019)

Conclusion

This chapter has presented statistical methods for computing and data science as an integrated whole, organized around three pillars: foundations, algorithms, and applications. Section 2 established the theoretical bedrock probability, statistical inference, Bayesian reasoning, information theory, and computational considerations—that gives statistical methods their logical coherence. Section 3 surveyed the core methods that form the working toolkit of data science: regression and generalized linear models, dimensionality reduction, classification and clustering, resampling and validation, regularization, nonparametric techniques, and time series analysis. Section 4 examined the computational machinery that turns these methods into working code: optimization, Monte Carlo and variational inference, numerical linear algebra, scaling strategies, and reproducibility practices. Section 5 demonstrated how methods and algorithms combine across healthcare, finance, natural language processing, computer vision, recommender systems, and network analysis. Section 6 addressed the challenges and emerging directions shaping the field's future, including fairness, uncertainty quantification, causality, privacy, and foundation models.

The central message is that statistics and computing are not separate disciplines but inseparable partners. Statistical theory provides the logic by which data yield trustworthy knowledge, while computation provides the means to apply that logic at scale. Neither is sufficient alone: theory

without implementation remains abstract, and implementation without theory risks producing confident but unreliable results. Mastery of both, and of the connections between them, is what distinguishes rigorous data science from mere data processing.

Several practical guidelines emerge from the preceding sections and are worth restating concisely. First, match the method to the question: predictive tasks, inferential questions, and structural discovery call for different tools, and the workflow in Figure 3.1 provides a starting point. Second, check assumptions: every method carries conditions independence, distributional form, linearity that must be verified rather than assumed. Third, quantify uncertainty: point estimates alone are rarely sufficient for good decisions, and intervals, posterior distributions, or conformal guarantees should accompany predictions wherever possible. Fourth, validate rigorously: cross-validation, bootstrap, and held-out test sets guard against overfitting and give honest estimates of generalization. Fifth, prioritize reproducibility: version control, dependency management, seeding, and literate analysis ensure that results can be verified and extended. Sixth, consider ethics and fairness from the outset, not as an afterthought, particularly in high-stakes domains where models affect people's lives. Finally, stay current: the field evolves rapidly, and methods that were impractical a decade ago are now routine, while new challenges continually emerge. Statistical methods have always been about extracting signal from noise, and that fundamental challenge has only grown more important as data has proliferated. The tools have evolved from hand-computed estimates to distributed inference on billions of records but the principles endure: model your uncertainty, validate your claims, and never confuse correlation with causation. This chapter has aimed to provide both a conceptual map and a practical toolkit for applying those principles in computing and data science. If it equips readers to approach data-driven problems with rigor, humility, and curiosity, it will have served its purpose.

References

1. Abdulah, S., Salvaña, M. L. O., Sun, Y., Keyes, D. E., & Genton, M. G. (2025). High-performance statistical computing (HPSC): Challenges, opportunities, and future directions. *WIREs Computational Statistics*, 17(4), e70052.
2. Aldridge, I., & Avellaneda, M. (2021). *Big data science in finance*. John Wiley & Sons.
3. Arabie, P., Hubert, L., & De Soete, G. (1996). *Clustering and classification*. World Scientific.
4. Ash, R. B. (1990). *Information theory*. Courier Corporation.
5. Botev, Z., Kroese, D. P., Taimre, T., & Vaisman, R. (2019). *Data science and machine learning: Mathematical and statistical methods*. CRC Press.
6. Bouveyron, C., Celeux, G., Murphy, T. B., & Raftery, A. E. (2019). *Model-based clustering and classification for data science: With applications in R* (1st ed.). Cambridge University Press. <https://doi.org/10.1017/9781108644181>

7. Brockwell, P. J., & Davis, R. A. (2009). *Time series: Theory and methods*. Springer Science & Business Media.
8. Cao, L., Yang, Q., & Yu, P. S. (2021). Data science and AI in FinTech: An overview. *International Journal of Data Science and Analytics*, 12(2), 81–99.
9. Casella, G., & Berger, R. (2024). *Statistical inference*. CRC Press.
10. Cover, T. M., Thomas, J. A., & Kieffer, J. (1991). *Elements of information theory* (Vol. 2). Wiley, New York.
11. Everitt, B. (2012). *Introduction to optimization methods and their application in statistics*.
12. Fan, J., Li, R., Zhang, C.-H., & Zou, H. (2020). *Statistical foundations of data science* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9780429096280>
13. Fodor, I. K. (2002). *A survey of dimension reduction techniques*. Lawrence Livermore National Lab., CA (US).
14. Gabriel, P., & Marco, C. (2019). Computational optimal transport with applications to data sciences. *Foundations and Trends® in Machine Learning*, 11(5–6), 355–607.
15. Gamerman, D., Lopes, H. F., & Gonçalves, F. B. (2026). *Markov chain Monte Carlo: Stochastic simulation for Bayesian inference*. CRC Press.
16. Gray, R. M. (2011). *Entropy and information theory*. Springer Science & Business Media.
17. Hammersley, J. (2013). *Monte Carlo methods*. Springer Science & Business Media.
18. Henley, S. S., Golden, R. M., & Kashner, T. M. (2020). Statistical modeling methods: Challenges and strategies. *Biostatistics & Epidemiology*, 4(1), 105–139.
19. Ibe, O. (2014). *Fundamentals of applied probability and random processes*. Academic Press.
20. Jacod, J., & Protter, P. (2004). *Probability essentials*. Springer Science & Business Media.
21. James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). Resampling methods. In *An introduction to statistical learning: With applications in Python* (pp. 201–228).
22. Jordan, M. I. (2013). On statistics, computation and scalability. *Bernoulli*, 19(4). <https://doi.org/10.3150/12-BEJSP17>
23. Kroese, D. P., & Rubinstein, R. Y. (2012). Monte Carlo methods. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(1), 48–58.
24. Lane, H., & Dyschel, M. (2025). *Natural language processing in action*. Simon and Schuster.
25. Lavine, B. K., & Mirjankar, N. (2000). Clustering and classification of analytical data. *Encyclopedia of Analytical Chemistry*, 11, 9689–9710.
26. Little, M. A. (2019). *Machine learning for signal processing: Data science, algorithms, and computational statistics*. Oxford University Press.
27. Mirkin, B. (2013). *Mathematical classification and clustering*. Springer Science & Business Media.

28. Mishra, S. K., Jeba Praba, J., & George, J. J. (2024). An emerging trend of bioinformatics and big data analytics in healthcare. *Digital Transformation in Healthcare*, 5, 159.
29. Rényi, A. (2007). *Probability theory*. Courier Corporation.
30. Ricaud, B., Borgnat, P., Tremblay, N., Gonçalves, P., & Vandergheynst, P. (2019). Fourier could be a data scientist: From graph Fourier transform to signal processing on graphs. *Comptes Rendus. Physique*, 20(5), 474–488.
31. Rocklöv, J., & Gayle, A. A. (2021). Data science and machine learning: Mathematical and statistical methods. D. P. Kroese, Z. I. Botev, T. Taimre, and R. Vaisman. *International Journal of Epidemiology*, 49(6), 2094–2095. <https://doi.org/10.1093/ije/dyaa072>
32. Sijtsma, K., & Emons, W. H. M. (2010). Nonparametric statistical methods. In *International encyclopedia of education* (pp. 347–353). Elsevier.
33. Sohil, F., Sohali, M. U., & Shabbir, J. (2022). An introduction to statistical learning with applications in R: By Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, *Statistical Theory and Related Fields*, 6(1), 87–87.
34. Sykes, A. O. (1993). *An introduction to regression analysis*.
35. Tatsat, H., Puri, S., & Lookabaugh, B. (2020). *Machine learning and data science blueprints for finance*. O'Reilly Media.
36. Verma, R. K., Kishor, K., & Jha, S. K. (2024). Big data analytics in bioinformatics and healthcare. In *Applications of parallel data processing for biomedical imaging* (pp. 25–43). IGI Global Scientific Publishing.
37. Volodina, V., & Challenor, P. (2021a). The importance of uncertainty quantification in model reproducibility. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2197), rsta.2020.0071, 20200071. <https://doi.org/10.1098/rsta.2020.0071>
38. Volodina, V., & Challenor, P. (2021b). The importance of uncertainty quantification in model reproducibility. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 379(2197), rsta.2020.0071, 20200071. <https://doi.org/10.1098/rsta.2020.0071>
39. Wang, C., Chen, M.-H., Schifano, E., Wu, J., & Yan, J. (2016). Statistical methods and computing for big data. *Statistics and Its Interface*, 9(4), 399.
40. Wu, Y. (1999). Statistical learning theory. *Technometrics*, 41(4), 377–378. <https://doi.org/10.1080/00401706.1999.10485951>
41. Zeroual, I., & Lakhouaja, A. (2018). Data science in light of natural language processing: An overview. *Procedia Computer Science*, 127, 82–91.
42. Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67(2), 301–320.

DATA ANALYTICS FOR ANIMAL HEALTH: APPLICATIONS, CHALLENGES AND FUTURE PERSPECTIVES

Yogesh C. Bangar*, S. S. Dhaka, Sunil Kumar and Yash Raj

Department of Animal Genetics and Breeding,

Lala Lajpat Rai University of Veterinary and Animal Sciences (LUVAS), Hisar 125004 India

*Corresponding author E-mail: yogeshbangar07@gmail.com

Abstract

Data analytics is increasingly important in animal-health management by converting diverse data into useful information for disease prevention, surveillance and veterinary decision-making. Animal-health data are generated from clinical and laboratory records, behavioural and sensor monitoring, environmental measurements and population-level surveillance, with production and genomic data providing additional information. Statistical methods form the foundation of animal-health analytics and should be selected according to the outcome type, data distribution and study design. Machine learning and artificial intelligence complement statistical approaches, particularly for complex sensor, image and video data. This chapter covers the major sources of animal-health data, data analytics workflow, outcome-specific statistical approaches, applications, challenges and future perspectives. Particular emphasis is given on regression, mixed-effects and survival models for analysing different health outcomes. The chapter also discusses data quality, model validation and generalizability. Overall, integration of reliable data, appropriate statistical methods, advanced computational approaches and veterinary knowledge can strengthen disease detection, surveillance, animal welfare and evidence-based health management.

Keywords: Animal Health, Data Analytics, Statistics, Machine Learning, Precision Farming.

1. Introduction

Animal health is fundamental to animal welfare, food security, farm profitability and sustainable animal production. Disease and other ailments can cause mortality, treatment costs, impaired welfare and reductions in productive and reproductive performance. Maintaining good health is therefore a major priority for animals managed under intensive farms as well as pasture-based and field conditions. Performance indicators such as growth, milk yield, reproduction and feed intake can also provide useful information for assessing changes in animal health.

Traditionally, animal health is evaluated through clinical examination, farmer or handler observation, laboratory testing and routine health records. Although these approaches remain essential, some health problems are recognized only after clinical signs become apparent. Data analytics provides a systematic framework for organizing, analysing and interpreting the

increasing volume of clinical, laboratory, behavioural, production, reproductive and environmental data generated from animals. Appropriate analysis of such data can identify disease patterns, quantify risk factors and support evidence-based health decisions (Niloofar *et al.*, 2021).

Statistical methods form the foundation of animal-health data analytics. Descriptive statistics can summarize disease frequency, mortality, treatment outcomes and health-related performance indicators, while inferential methods allow relationships between health outcomes and potential risk factors to be evaluated. Regression models, mixed-effects models, survival analysis and longitudinal-data methods are particularly important because animal-health datasets commonly contain repeated observations, clustered data and time-dependent outcomes. Statistical analysis therefore enables estimation of disease risk, identification of important determinants and evaluation of interventions while accounting for biological and management-related variation (Petrie and Watson, 2013).

Precision livestock farming and sensor technologies are further increasing the availability of longitudinal animal-health data (Kleen and Guatteo, 2023; Curti *et al.*, 2023). Machine learning and artificial intelligence can complement conventional statistical approaches when analysing large or complex datasets (García *et al.*, 2020), but their outputs require appropriate validation and biological interpretation. Therefore, appropriate statistical analysis, together with advanced computational methods when needed, can turn animal-health data into useful information for disease prevention, surveillance and improved animal welfare.

2. Sources of Data for Animal-Health Analytics

Animal-health analytics uses data from multiple sources to assess health status, identify disease risk factors and monitor changes in individual animals, herds and populations. Integration of these data provides a stronger basis for statistical analysis and evidence-based veterinary decision-making (Niloofar *et al.*, 2021).

2.1 Clinical and Laboratory Data

Clinical records are fundamental sources of animal-health information and include clinical signs, disease diagnosis, body temperature, vaccination, treatment, medication, recovery and mortality. Laboratory data include haematology, serum biochemistry, microbiology, parasitology, serology, pathology and molecular tests. Diagnostic imaging, including radiography, ultrasonography and thermal imaging, provides additional objective information for disease assessment.

2.2 Behavioural and Sensor Data

Changes in behaviour can provide early indications of disease, stress and impaired welfare. Wearable sensors continuously measure activity, feeding, rumination, lying, standing and movement, while cameras and other non-contact technologies enable continuous monitoring

without any disturbances to animals. Repeated measurements help identify changes from an animal's normal health and behaviour.

2.3 Production and Reproductive Data

Various records such as milk yield, body weight, growth, feed intake, fertility and other reproduction parameters provide supplementary information on animal health. Changes in these variables may accompany disease or physiological stress and can therefore be included as health-related outcomes or predictors in statistical analyses.

2.4 Environmental and Genomic Data

Environmental parameters such as temperature, humidity, air quality, nutrition and water availability generally influence disease occurrence and welfare. Genomic data additionally provide information on genetic variation, disease susceptibility and resistance (Berry, 2023).

2.5 Herd- and Population-Level Data

Disease reports, outbreak records, animal movements, mortality and antimicrobial-use records provide important epidemiological information. Their analysis help in identifying temporal and spatial disease patterns, high-risk populations and emerging health problems, thereby supporting disease surveillance and veterinary public-health decisions (Guitian *et al.*, 2023).

Overall, accurate, complete and consistently recorded data from multiple sources provide the foundation for reliable statistical analysis of animal health.

3. Data Analytics Workflow

Data analytics should begin with a clear biological or animal-health question. The objective may be to identify risk factors for mastitis, detect lameness at an early stage, evaluate treatment outcomes or predict mortality. A clearly defined objective helps determine which animals, variables and time periods should be included in the analysis. Relevant data are then collected from clinical records, laboratory tests, farm records, sensors or other available sources. Accurate animal identification and recording dates are essential for linking information correctly.

Data cleaning is an important step before statistical analysis. Animal-health datasets may contain missing values, duplicate records, incorrect dates, inconsistent units, recording errors or biologically unrealistic observations. These problems should be identified and carefully examined. Variable names, units and categories should also be standardized. Missing values and outliers should not be removed automatically because they may contain useful biological information or indicate problems in data collection.

After cleaning, the data are organized and suitable variables are prepared for analysis. For example, repeated activity measurements may be summarized as daily averages or changes from an individual animal's normal level. Descriptive statistics and graphical methods can first be used to understand data distribution, disease frequency and general patterns. Appropriate

statistical methods are then selected according to the research question, type of outcome and structure of the data.

The results should be evaluated for statistical reliability and biological relevance. Where predictive models are used, validation with independent animals, herds, farms or time periods is important to determine their practical usefulness. Finally, analytical findings should be translated into meaningful actions, such as examining high-risk animals, conducting diagnostic tests or increasing disease surveillance. Thus, effective data analytics combines good-quality data, appropriate statistical methods and biological and veterinary knowledge to support reliable animal-health decisions (Niloofar *et al.*, 2021).

4. Analytical Approaches

The choice of analytical method in animal-health studies mainly depends on the type and distribution of the outcome variable, study design and structure of the data. Animal-health outcomes may be continuous, binary, categorical, count, ordinal, repeated or time-to-event (Petrie and Watson, 2013). Therefore, selecting a statistical method appropriate to the outcome is essential for obtaining reliable and biologically meaningful results.

4.1 Continuous Outcomes

Continuous outcomes can take a range of numerical values e.g. body temperature, body weight, milk yield, blood parameters and growth rate. Linear regression is commonly used to examine the relationship between a continuous outcome and one or more explanatory variables. For example, linear regression can assess whether body temperature is associated with age, season or disease status. Analysis of variance (ANOVA) can compare mean values among groups, while linear mixed-effects models are appropriate when measurements are repeatedly collected from the same animals or animals are clustered within farms.

4.2 Binary Outcomes

Binary outcomes have two possible categories, such as diseased/healthy, infected/non-infected, or dead/alive. Logistic regression is widely used for these outcomes. It estimates the probability of an event and determines its association with potential risk factors. For example, logistic regression can evaluate whether parity, season, previous mastitis and management practices are associated with the probability of mastitis. Results are commonly expressed as odds ratios with confidence intervals.

4.3 Multicategory Outcomes

Some health outcomes contain more than two categories without a natural order, such as healthy, respiratory disease, digestive disease and other disease. Multinomial logistic regression can be used for such outcomes. It allows the probability of each disease category to be compared with a selected reference category and can identify factors associated with different disease conditions.

4.4 Ordinal Outcomes

Ordinal outcomes consist of categories with a meaningful order, such as lameness scores, body-condition scores or disease-severity categories (mild, moderate and severe). Ordinal logistic regression is suitable for analysing these outcomes because it retains the natural ordering of the categories. For example, it can evaluate factors associated with increasing severity of lameness.

4.5 Count Outcomes

Count outcomes represent the number of events occurring within a specified period, such as number of disease episodes, veterinary treatments, parasite counts or clinical cases. Poisson regression is commonly used for count data. However, animal-health data often show greater variability than expected under the Poisson model. In such situations, negative-binomial regression is generally more appropriate. Rates can also be analysed by incorporating the population or time at risk into the model.

4.6 Proportion and Prevalence Outcomes

Animal-health studies frequently examine proportions, such as the proportion of infected animals, treatment success or herd-level disease prevalence. Simple proportions can be presented with confidence intervals, while comparisons between groups may initially use chi-square or Fisher's exact tests. For multivariable analysis of individual binary health status, logistic regression is commonly used. When the objective is to directly estimate prevalence ratios or risk ratios, appropriate generalized linear models may be considered.

4.7 Repeated and Clustered Outcomes

Animal-health datasets commonly contain repeated measurements from the same animal or observations from animals grouped within herds, farms, villages or regions. Such observations are not statistically independent. Mixed-effects models and generalized linear mixed models can account for this clustering. For example, repeated disease status from cows within several farms can be analysed by including animal and/or farm effects in the model. These methods are particularly important for longitudinal and multicentre animal-health studies.

4.8 Time-to-Event Outcomes

When the outcome concerns the time until an event occurs, survival analysis is appropriate. Examples include time to disease onset, recovery, recurrence, culling or death. Kaplan–Meier analysis describes survival or event-free probabilities over time, whereas the Cox proportional-hazards model evaluates the association between explanatory variables and the rate of an event. Survival analysis is particularly useful because animals that have not experienced the event by the end of follow-up can still contribute information.

4.9 Predictive Outcomes and Machine Learning

Statistical models such as logistic regression, mixed models and survival models can also be used for prediction. When datasets are large and contain numerous or complex predictors,

machine-learning methods such as decision trees, random forests and gradient boosting may provide complementary approaches (García *et al.*, 2020). Artificial intelligence and deep learning are particularly relevant to specialised data such as images, videos and high-frequency sensor measurements. Computer vision, for example, has been used to assess feeding behaviour and detect animal abnormalities (Bezen *et al.*, 2020; Bhujel *et al.*, 2025). However, these approaches require adequate datasets and careful external validation because generalizability remains an important concern (Rahman *et al.*, 2025).

4.10 Selection and Evaluation of the Analytical Method

The analytical method should therefore be selected according to the outcome type, data distribution, study design, repeated or clustered structure and analytical objective. Model assumptions should be examined before interpretation. Effect estimates, confidence intervals and biological importance should be reported alongside statistical significance. For predictive disease models, sensitivity, specificity, positive predictive value and area under the receiver-operating-characteristic curve (AUC) are useful measures, whereas mean absolute error (MAE) and root mean square error (RMSE) can evaluate predictions of continuous outcomes. Appropriate statistical analysis ensures that animal-health data are converted into reliable and practically meaningful information.

5. Challenges and Limitations

The reliability of animal-health analytics depends strongly on data quality, completeness and consistency. Missing observations, incorrect animal identification, inconsistent disease definitions, recording errors and differences in measurement methods can introduce bias and affect statistical results. Integration of clinical, laboratory, behavioural, environmental and production data may also be difficult because these data often originate from different systems and time periods.

Animal-health datasets frequently have complex structures, including repeated observations, clustering of animals within herds or farms, unequal follow-up periods and relatively few disease cases. Inappropriate statistical methods or failure to consider confounding, missing data and model assumptions can produce misleading conclusions. Generalizability is another challenge because results obtained from one breed, herd or production system may not apply to other populations. Machine-learning and AI approaches additionally require sufficient data and careful validation to avoid overfitting and poor interpretability (Bhujel *et al.*, 2025). Therefore, analytical findings should always be interpreted with biological and veterinary knowledge.

6. Future Perspectives

The future of animal-health analytics will increasingly involve AI and advanced digital monitoring technologies. Integration of clinical, laboratory, behavioural, environmental and sensor data will enable AI models to identify complex disease patterns and provide earlier health

warnings. Deep learning and computer vision may improve automated detection of lameness, abnormal behaviour and other health conditions from images, videos and sensor data. Statistical methods will remain essential for evaluating risk factors, longitudinal trends and model validity. Future progress will require larger standardized datasets, external validation and interpretable AI models that can provide reliable decision support for veterinarians and improve preventive health management and animal welfare.

Conclusions

Data analytics provides a systematic approach for converting diverse animal-health data into useful information for disease prevention, surveillance and veterinary decision-making. Clinical, laboratory, behavioural, sensor, environmental and population-level data form the major sources, while production and genomic data provide supporting useful information. Reliable analysis requires accurate data collection, careful cleaning and appropriate integration. Statistical methods remain central and should be selected according to the outcome type and data structure, including regression, mixed-effects and survival models. Machine learning and AI can complement these methods for complex datasets. Combining reliable data, appropriate statistical analysis and veterinary knowledge can improve animal health, welfare and management.

References

1. Berry, D. P. (2023). Large-scale phenotyping and genotyping: State of the art and emerging challenges. In I. Kyriazakis (Ed.), *Smart livestock nutrition*. Springer.
2. Bezen, R., Edan, Y., & Halachmi, I. (2020). Computer vision system for measuring individual cow feed intake using RGB-D camera and deep learning algorithms. *Computers and Electronics in Agriculture*, 172, 105345.
3. Bhujel, A., Wang, Y., Lu, Y., Morris, D., & Dangol, M. (2025). A systematic survey of public computer vision datasets for precision livestock farming. *Computers and Electronics in Agriculture*, 229, 109718.
4. Curti, P. F., Selli, A., Pinto, D. L., Merlos-Ruiz, A., Balieiro, J. C. C., & Ventura, R. V. (2023). Applications of livestock monitoring devices and machine learning algorithms in animal production and reproduction: An overview. *Animal Reproduction*, 20(2), e20230077.
5. García, R., Aguilar, J., Toro, M., Pinto, A., & Rodríguez, P. (2020). A systematic literature review on the use of machine learning in precision livestock farming. *Computers and Electronics in Agriculture*, 179, 105826.
6. Guitian, J., Arnold, M., Chang, Y., & Snary, E. L. (2023). Applications of machine learning in animal and veterinary public health surveillance. *Scientific and Technical Review*, 41(2).

7. Kleen, J. L., & Guatteo, R. (2023). Precision livestock farming: What does it contain and what are the perspectives? *Animals*, 13(5), 779.
8. Niloofar, P., Francis, D. P., Lazarova-Molnar, S., Vulpe, A., Vochin, M.-C., Suci, G., Balanescu, M., Anestis, V., & Bartzanas, T. (2021). Data-driven decision support in livestock farming for improved animal health, welfare and greenhouse gas emissions: Overview and challenges. *Computers and Electronics in Agriculture*, 190, 106406.
9. Petrie, A., & Watson, P. (2013). *Statistics for veterinary and animal science* (3rd ed.). Wiley-Blackwell, Oxford, UK.
10. Rahman, Z. U., Mohd Asaari, M. S., & Ibrahim, H. (2025). Advancing animal farming with deep learning: A systematic review. *Computers and Electronics in Agriculture*, 237, 110674.

ARTIFICIAL INTELLIGENCE AND ORGANIZATIONAL TRANSFORMATION: A HUMANISTIC APPROACH TO CONTEMPORARY MANAGEMENT

Ranjit Kumar Paswan

Department of Commerce,

Kazi Nazrul University, Asansol, West Bengal, India

Corresponding author E-mail: ranjit.paswan81@gmail.com

Abstract

Artificial Intelligence is considerably altering the organizational activities by changing the several activities such as management procedures, leadership styles, work culture and decision-making process. The inclusion of AI provides an opportunity for organization to improve operational efficiency, innovation and productivity. However, integration of AI requires assessment of how organization manages people. As AI arrangements have huge effect on the organizational functions, it is vital to consider human dignity, welfare of employee, transparency and autonomy. This chapter seeks to assess the relevance of humanistic approach to management development with the help of AI. The successful human-AI integration depends on advancing technological competencies, employees' participation and learning process. Ultimately, the technological advancement alone should not be considered for success of AI-enabled business transformation but human contribution, ethical governance and sustainable development of organization should also be considered.

Keywords: Artificial Intelligence, Humanistic Approach, Organizational Transformation, AI-Human Integration, Leadership.

1. Introduction

Today's world is surrounded by computer system. Artificial Intelligence is one of the emerging proficiencies of computer system or machine which is ready to perform any task with the help of human intelligence. In modern corporate atmosphere, AI has become one the most transformative technical extension. Rapid change in machine learning, robotics and data analytics have increasingly extended the organizational capacity in collecting, processing and utilizing information. The earlier technological revolutions were primarily elaborate routine movements. However, AI systems taking part in solving complex task, forecasting, strategic planning and decision-making. Consequently, AI becomes a key driver for organizational transformation. **Najana (2024)** advocates that inclusion of AI in organizational process has changed the operational and strategic paradigm of global business. Further, the author added that AI is significantly restructuring organizational process particularly in the area of financial services, customer satisfaction and product development. The developmental phase of the inclusion of AI in business operation may involve in deep core workflows. The business

organization may not reap benefits of AI from standalone workflow rather deep implantation can generate value of the organizations.

According to World Economic Forum “Organizations moved rapidly from experimentation to capability, advancing through pilots, proofs of concept and early developments”. The increasing acceptance of AI is continuously remodeling the processes, structure, managerial systems of organizations. Organizations are integrating AI into its various activities such as human resource management, finance, marketing, customer relationship management, supply-chain management etc. It is contended that there is a multifaceted condition in controlling a large data set. With the implementation of AI, the arrangements can expand operational proficiency, certify speedy decision-making, diminish working costs and empowers organization to acquire innovative creations. AI enhances decision-making by processing the large set of data, identifies pattern and providing data-driven knowledge (IBM, 2025). AI-driven changes encompass outside technological implementation as it leads to improve working environment, managerial competencies, employee upgradation, organizational liaisons and the trend of value creation.

Despite these possible advantages, the increasing adoption of AI into organizational activities have raised human, ethics and social issues. The automation task and growing dependence on decisions based on algorithm may have negative influence on employment, privacy, discrimination, accountability and transparency. It may threaten the individual autonomy in business decision-making and reduce human judgement. Therefore, the outcomes-built questions due to integration of AI which cannot be understood exclusively through better financial health. The other important factor like human values must also be considered apart from technological transformation. As the organizations are considered as social institutions in which technology is used to relate people, values, culture and organizational structures. Accordingly, AI implementation needs to establish balance between technological efficiency and human value. A humanistic approach in organizational practice considers human as valued individual empowering them with autonomy and also focusing on human dignity, personal flourishing, safety and ethical responsibility. Therefore, harmonization of AI-driven changes and human competence becomes essential to redesign. These organizational changes may create a business environment with responsible and sustainable organizational growth. The growing incorporation of human judgement and AI-based decision-making highlights the need for humanistic approach. AI may provide refined analysis and information but managerial decisions may often include appropriate understanding, judgement based on ethical ground considering stakeholders interests. In light of these broader perspectives, the present chapter evaluates the association between artificial intelligence and change in organization structures from a humanistic point of view. It studies how AI is redefining operational activities of an organization, their managerial roles, work cultures and integrated role of human and artificial intelligence touching moving along with ethical relation and social challenges. Additionally, the chapter also explores how

management with humanistic approach enable AI-driven developments towards a more human-oriented, ethical, comprehensive and trustworthy approach. Finally, the chapter recommends a conceptual framework of integration of technological expansion and human values to accelerate sustainable and reliable organizational makeover in the era of artificial intelligence.

2. Objectives

- To know the role of artificial intelligence in organizational renovation in contemporary business environment.
- To examine the acceptance of AI-driven transformation for managerial tasks and human-AI cooperation.
- To examine the ethical, social and organizational challenges connected to growing use of AI in organizations
- To ascertain the role of humanistic management beliefs in directing AI-based organizational change.
- To propose a theoretical outline for attaining ethical, responsible and human-centered organizational renovation with artificial intelligence.

3. Emerging Role of Artificial Intelligence in Organization

a. Evolution of Artificial Intelligence

Earlier a strong discussion has been done on the economic impacts of AI, where digital technologies like AI and big data disorder business models, increase efficiency, reduce waste and empower firms to become active to improve stakeholder experience (Chauhan *et al.*, 2022). In the words of Dr Alan Litchfield “People believe that technology is something that is different from the person. Actually, technology is a demonstration of our existing state of mind. Technology has often supported what we have been doing or considering.” The impression of artificial intelligence occurred long before but came to formal platform in 1956. Previously, intelligence researchers assumed intelligence had to be cautiously invented by humans. Darwin’s theory of evolution exposed that complex and flexible systems can change naturally through evolution without being constructed by humans. The movement contends that AI researchers should likewise identify the probability of developed machine intelligence, rather than centering only on human-designed orders (Spector, 2006).

b. Key Concept of Technological Development

Artificial Intelligence incorporates a set of technologies which allow computer systems to accomplish projects that conventionally need human intelligence such as learning, arguing, forecasting, decision-making and content creating. In recent decades, AI has improved into a broader technical network which includes generative artificial, machine learning and intelligence automation. Even these concepts are interrelated, each of them has a distinct role and uses within organization. These technologies may work together in an organization. For example, if an organization wants to forecast demand, it may use machine learning. Further, generative AI may

be used to review the forecast and intelligent automation may help in instigate inventory-related plans. This integration shows how AI can combine the organizational activities showing its analytical capacities.

4. Application of AI in Organization

Organizations of various diversified industries are now integrating AI mechanisms into their work culture aiming at increasing operational efficiency, productivity and invention (Benbya *et al.*, 2020). However, implementation of AI causes significant challenges particularly relating to its effects on workplace and organizational culture (Dwivedi *et al.*, 2021). Furthermore, AI-enabled programs may boost communication and integration among employees, which may lead to encourage their participation in decision-making and supporting empowered culture (Dasgupta and Wendler, 2019).

5. Humanistic Approach to Modern Management

The humanistic approach to modern management considers human beings as the main focus area of organizational functions and decision-making. On the contrary traditional management system focuses mainly on efficiency, productivity and overall performance of organizations. Humanistic management considers employees as individual providing dignity and autonomy with social needs and ambitions. Therefore, the modern approach pursues equalization in organizational purposes and human welfare, ethical concern and social respect. The humanistic approach incorporates a philosophical perception which recognizes the importance of human values within the organizational employment (Diawati *et al.*, 2023). With the progress of technology, it is relevant to assert that in modern management the inclusion of artificial intelligence becomes appropriate in several organizational functions such as decision-making, job alteration and workplace connection. The humanistic approach plays an important role in creating functional and productive environment. It considers the human values, employee wellbeing and innovative ideas of human being in workplace. With this application an organization may create a supporting business environment by encouraging personal development, increasing leadership and employee satisfaction (Asdlori, 2023). Although, artificial intelligence may improve productivity, upgrading and methodical capacities along with job security, employee freedom, inequality etc. A humanistic approach therefore, advocates that technological enhancement may accompany human capabilities. Within an organization several sources form the bases of humanistic management. Human dignity comprises of value of every individual within an organization. Respect for individual contains acknowledgement of their knowledge and competencies.

AI is playing an important role in moving forward the organizational productivity. Businesses are adopting AI to update their supply chain, reduce costs and increase client satisfaction. AI structures can oversee stock management and trace distributions in real time (Govindharaj, 2025). Arranging and automating human resource management practice with AI can boost

organizational effectiveness by minimizing managerial inactivity and ensuring technical reliability. However, this benefit mainly narrates data entry workflows and payroll processing where constant diligence is required (Yang, 2026).

Collectively, these forms give a human-centric bases for AI-enabled organizational change. They propose that success of an organization should be examined not only through technological and financial health but also through the scope to which transformation upholds dignity, welfare, transparency, participation and balanced human development.

6. AI-Human Collaboration and Change in Working Environment

The collaboration of artificial intelligence into organizations is changing tasks, workplace relationships, expertise, and work culture. AI-human collaboration integrates the various activities such as data handling, analytical capacity and automation strength with human ideas, ethical decisions and innovative thinking. AI may regularize the abilities of employees and enable them to involve in multifaceted and advanced activities. Human-AI association offers opportunities to expand work competence and enhance decision quality. At the same time, it establishes challenges regarding technology implementation, reconstruction and redefining work (Parker and Grote, 2022). In collaborative work situations, understanding how employees actively familiarise and reform their work with such AI-enabled surrounding and how these adoptions develop job performance, has become an imperative area of investigation in organizational behaviour (Bao *et al.*, 2026). AI adoption also enhances the need for constant learning, digital knowledge, compliance and reskilling. Apart from the benefit of AI-human collaboration, it gives challenges towards job and data insecurity, falling employee autonomy and excessive workplace censoring. Depending more on automated techniques may lead to weak interpersonal relationships and employee confidence if they are neglected.

From a human-centred point of view, AI-Human combination may become effective if the organizations respect employee's dignity, wellbeing, and professional development. Impartial communication and transparent organizational conduct are essential for establishing trust and sense of responsibility among employees. For achieving operational effectiveness and human wellbeing a significant association need to be develop among technological advancement and human capabilities.

7. Ethical and Social Challenges AI-driven Organizational Transformation

Unlike traditional information systems, AI systems acquire the ability to study from data, make forecast and figure human decision-making. This ability may lead to ethical worries which may include transparency, privacy, fairness and responsibility. Corporates usually depend on AI solutions shaped by external parties or using historical data that may incorporate biases or mistakes. Consequently, AI-based decisions may accidentally lead to discrimination or destructive outcomes. AI technologies provide significant operational benefits; their integration also constructs ethical challenges that must be look into and responsible use can be promoted. As

AI systems depends on large dataset, the data privacy may become a serious concern as it holds sensible personal data. The collection, preservation and usage of such data can enhance the risk of mishandling and compromising privacy of individuals (Zuboff, 2019 as cited in Elson 2024,). The advancement of technology needs to be considered carefully. As AI becomes increasingly important across different sectors, organizations must safeguard its integration by establishing transparent and fair principles which defines accountability of its implementation. It requires strict governance frameworks for greater clarity in decision making and identifying algorithmic bias (Elson, 2024). The role of AI in concentrating global challenges is likely to improve its potential for social and organizational change. In long run, growth of AI will depend on establishment of ethical and transparent governance. Further, it is argued that significant role of accountable artificial intelligence may enhance the social and ethical governance and also increase value-growth in different sectors. The effective public policy, fairness and unambiguity are important for confirming the meaningful contribution of AI towards sustainable social and economic development (Ramya, 2026).

8. Leadership and Management in the Era of AI

Artificial intelligence has carved a position in transforming leadership and management style in areas such as strategic planning, decision making and managerial practices. AI-enabled resources provide leaders with enhance data-driven knowledge and real time information which supports the advanced decision making and operational efficiency. Managers also needed to incorporate leadership approach, competent digital practice and create an innovative culture with continuous learning in AI-enabled business environment. For humanistic management perspective, leaders must confirm AI implementation with responsible method and after considering employee wellbeing, transparency, fairness and accountability. Therefore, leadership in the era of AI encompasses harmonizing technical abilities with human decision, ethical responsibility and broader managerial practices. Leadership practice experiencing a crucial transformation with the influence of AI, changing leader's role from traditional processes to crossing emerging and undefined challenges. This requires refinement in traditional practice by balancing data-driven knowledge with human brain. Leadership in AI era denotes evolution towards a more integrated, futuristic and ethically enabled organizational leadership (**Pandey, 2025**). The incorporation of Generative Artificial Intelligence (GenAI) into organization is reforming the leadership competencies. In this challenging business environment, the leaders must enhance their skills to manage organizational changes impacted by AI (Gomes *et al.*, 2026).

Conclusions

Artificial Intelligence is playing a significant role in transforming organizational management practices, leadership skills, decision-making strategies and work culture. The implication of AI allows the organization to enhance operational efficiency, productivity and innovation. However, the implementation of AI needs a careful consideration of people within the organization. As

inclusion of AI becomes more important in organizational activities, the other issues such as employee wellbeing, dignity, transparency and accountability must not be ignored.

This chapter has highlighted how humanistic approach is relevant to AI-driven organizational change. Although, the capability of AI is much more inclusive, its integration must not compromise wellness and dignity of employee. A humanistic approach gives an understanding about the compliance of human values, fairness, responsibility with AI adoption. Finally, the successful organizational transformation should not be determined only by advancement of technology and productivity but also by human contribution, ethical governance and sustainable development of organization.

References

1. Asdlori. (2023). The importance of humanistic approach in human resource management education. *Indo-MatchEdu Intellectuals Journal*, 397–419.
2. Bao, J., Wang, Q., Pan, Y., Wang, X., & Zheng, Y. (2026). Crafting work with AI: Human-AI collaborative job crafting, human-AI fit and employee job performance. *Frontiers in Psychology*, 1–13.
3. Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Artificial intelligence in organizations: Current state and future opportunities. *MIS Quarterly Executive*.
4. Chauhan, C., Parida, V., & Dheer, A. (2022). Linking circular economy and digitalisation technologies: A systematic literature review of past achievements and future promises. *Technological Forecasting and Social Change*, 1–18.
5. Diawati, P., Gadzali, S. S., Aziz, M. N., Ausat, A. A., & Suherlan, S. (2023). The role of information technology in improving the efficiency and productivity of human resources in the workplace. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 296–302.
6. Elson, A. (2024, November). *AI-driven transformation in industry: Balancing innovation with ethical and operational challenges*. <https://doi.org/10.13140/RG.2.2.34402.13766>
7. Forum, W. E. (2026). *Organizational transformation in the age of AI: How organizations maximize AI's potentia*. World Economic Forum.
8. Gomes, J. L., Nascimento, L. A., & Vilela, K. C. (2026). The rise of leadership skills in the era of generative artificial intelligence: A systematic review. *Digital Education Review*, 165–186.
9. Govindharaj, D. (2025). The role of AI in modern management. *ComFin Research*, 192–197.
10. IBM. (2025, December 15). *Take advantage of AI and use it to make your business better*.
11. Asdlori. (2023). The importance of humanistic approach in human resource management education. *Indo-MatchEdu Intellectuals Journal*, 397–419.

12. Bao, J., Wang, Q., Pan, Y., Wang, X., & Zheng, Y. (2026). Crafting work with AI: Human-AI collaborative job crafting, human-AI fit and employee job performance. *Frontiers in Psychology*, 1–13.
13. Benbya, H., Davenport, T. H., & Pachidi, S. (2020). Artificial intelligence in organizations: Current state and future opportunities. *MIS Quarterly Executive*.
14. Chauhan, C., Parida, V., & Dheer, A. (2022). Linking circular economy and digitalisation technologies: A systematic literature review of past achievements and future promises. *Technological Forecasting and Social Change*, 1–18.
15. Diawati, P., Gadzali, S. S., Aziz, M. N., Ausat, A. A., & Suherlan, S. (2023). The role of information technology in improving the efficiency and productivity of human resources in the workplace. *Jurnal Teknologi Dan Sistem Informasi Bisnis*, 296–302.
16. Elson, A. (2024, November). *AI-driven transformation in industry: Balancing innovation with ethical and operational challenges*. <https://doi.org/10.13140/RG.2.2.34402.13766>
17. Forum, W. E. (2026). *Organizational transformation in the age of AI: How organizations maximize AI's potentia*. World Economic Forum.
18. Gomes, J. L., Nascimento, L. A., & Vilela, K. C. (2026). The rise of leadership skills in the era of generative artificial intelligence: A systematic review. *Digital Education Review*, 165–186.
19. Govindharaj, D. (2025). The role of AI in modern management. *ComFin Research*, 192–197.
20. IBM. (2025, December 15). *Take advantage of AI and use it to make your business better*.
21. Litchfield, D. A. (2021, September 24). *The evolution of artificial intelligence, part 1: A double-edged sword*.
22. Najana, M., Bhattacharya, S., Kewalramanu, C., & Pandiya, D. K. (2024). AI and organizational transformation: Navigating the future. *International Journal of Global Innovations and Solutions*.
23. Pandey, V. (2025). Leadership in the AI era: Navigating and shaping the future of organizational guidance. *International Journal of Science and Research Archive*, 1737–1747.
24. Parker, S. K., & Grote, G. (2022). Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. *Applied Psychology*, 1171–1204.
25. Ramya, P. (2026). AI for sustainable social transformation: Ethics, governance and inclusive innovation. *Academe Journal of Education and Psychology*, 364–372.
26. Spector, L. (2006). Evolution of artificial intelligence. *Artificial Intelligence*, 1251–1253.
27. Yang, M. (2026). Human-centered AI adoption in organizational human resource management: Determinants and implementation strategies. *AEM JOURNAL.ORG*, 2723–2730.

AI-DRIVEN HUMAN RESOURCE MANAGEMENT AND EMPLOYEE WELL-BEING: AN INTERDISCIPLINARY PERSPECTIVE ON THE FUTURE OF WORK

V. T. Vidhya Poorani and S. Subramanian

Department of Management Studies,

Amet University, 135, East Coast Road, Kanathur, Chennai 603112

Corresponding author E-mail: vtvidhyapoorani@gmail.com, subramaianc2@gmail.com

Abstract

Artificial Intelligence (AI) is increasingly transforming Human Resource Management (HRM) by changing how organizations recruit, develop, evaluate, engage and retain employees. AI-enabled HR systems can automate repetitive activities, analyse large volumes of workforce data, support decision-making and provide personalized learning and performance feedback. However, the growing use of AI in people management also creates concerns related to employee autonomy, privacy, fairness, transparency, trust, job security and psychological well-being. This chapter examines the relationship between AI-driven HRM and employee well-being from an interdisciplinary perspective integrating HRM, organizational behaviour, psychology, information technology, ethics and sustainability. Drawing on recent literature, the chapter argues that AI does not have a uniformly positive or negative effect on employees. Instead, outcomes depend on how AI is designed, implemented, governed and interpreted by employees. Employee autonomy, perceived fairness, organizational trust and digital competence are proposed as important mechanisms through which AI-driven HRM may influence well-being. Human oversight, employee participation and ethical AI governance are further identified as important conditions for achieving positive outcomes. A conceptual framework is developed to explain these relationships and to provide directions for future empirical research. The chapter concludes that organizations should move from an automation-oriented approach towards a human-centred AI-HRM model in which technology complements rather than replaces human judgement.

Keywords: Artificial Intelligence, Human Resource Management, Employee Well-Being, Algorithmic Management, Employee Autonomy, AI Trust, Digital Workplace, Ethical AI.

1. Introduction

The advent of Artificial Intelligence (AI), machine learning, automation, big-data analytics, and digital platforms has rapidly transformed the workplace. Organizations have increasingly adopted AI to assist with recruiting, employee selection, performance management, learning, and employee development, workforce planning, and employee engagement, in addition to

managerial decision-making. Consequently, Human Resource Management (HRM), historically the domain of human judgment and interaction, has become an increasingly technology-assisted and data-driven profession.

The use of AI in HRM extends beyond technological change to a broader organizational evolution impacting employees, their managers, HR professionals and ultimately the nature of work itself.

Through a systematic review of 60 studies, Pereira *et al.* (2021) have shown that AI can affect work place outcomes at the different levels of analysis across various HRM functions. They highlighted the necessity to frame AI within a multi-stage, end-to-end process involving antecedents, technological phenomenon, and employee and organizational outcomes. Pessac *et al.* (2023) reviewed 67 peer-reviewed articles and developed a strategic framework of AI-assisted HRM; their work emphasizes how AI can generate employee and organizational outcomes. More recently, Gong *et al.* (2025) provided an overview of the state of AI-HRM research and discussed emerging research streams on AI-enabled collaboration, AI-driven workplaces, AI-enabled business models, and AI-powered innovation.

The trans-disciplinary nature of AI-HRM has also been underscored by Dima *et al.* (2024), who have analyzed 43 peer-reviewed articles and provided insights on the ways AI affects HR actions and the roles and tasks of HR professionals, line managers, and employees.

Their analysis has provided a clear indication that AI-HRM not only concerns the topic of technology HR but encompasses various related discipline. The potential benefits of using AI are clear: automation of many administrative tasks; processing of large datasets; accelerated HR decisions; personalized learning and support; and identification of patterns that are difficult to see with traditional HR methods. However, the introduction of AI also raises concerns about job losses, surveillance, and privacy, bias in algorithms, and diminished control and employee trust.

Given these concerns, an important aspect in understanding the future of AI-enabled work involves focusing on employee wellbeing. AI can be beneficial in that it can alleviate dull and repetitive tasks and improve working conditions by assisting employees to make smarter decisions. On the other hand, algorithmic performance management and automated decision-making may generate greater pressure at work and a decreased sense of control over work tasks. In line with these considerations, the central argument of this chapter is that the impact of AI-driven HRM on employee well-being hinges not merely on the technologies deployed but on how they are designed, implemented, and experienced by employee.

2. AI-Driven Human Resource Management

AI-driven HRM means applying artificial intelligence and associated technologies to facilitate, enhance, or automate HR tasks and employee-related judgments. Examples of AI systems include machine learning, natural language processing, predictive analysis, recommendation

system, and automatic decision support solutions. According to literature, there are various of AI applications in several HR functions. According to Bujold *et al.*'s (2024) systematic review of 107 empirical and scholarly papers, there are several HR applications using AI, including in talent acquisition, job performance management, talent management, work planning, health and well-being, and compensation, along with ethical, fair and responsible AI concerns such as fair AI (e.g. Absence of discrimination, fairness) (Bujold, Al-Bustami *et al.* 2024), fairness and trust, privacy, security, transparency and explainability, human oversight/participation etc.

2.1 AI in Recruitment and Selection

AI in Recruitment and Selection Recruitment is arguably the most visible area of AI in HRM. AI technologies can screen candidate profiles (or CVs), detect specific characteristics of candidate profile, screen candidates based on job role and aid recruiters in determining candidate ranking. The primary benefit gained is the increased efficiency; an organization which might receive thousands of applications will be able to shave resources and time on initial screening.

Chen (2022) researched about human-AI collaboration within recruitment; he pointed out how AI can play a role in mitigating human biases inherent in recruitment practices.

However, as any AI tool, there is no guarantee of fairness; an AI recruitment tool trained on biased historical data is going to repeat similar biases in practice. Data governance, fairness testing and human intervention is therefore crucial for AI aided recruitment. Fair-by-design approach has been suggested by researchers who explore issues of responsible AI for HRM technology. The approach argues not for enhancement of predictive accuracy but for testing and controlling of discriminatory results through development phase of algorithms (2022)

2.2 AI in Performance Management

Beyond objective KPIs, AI can assist in performance management by interpreting metrics of employee productivity, quality of work product, employee reviews and behavioral patterns. These tools are often intended to help manager's spot patterns in performance and deliver more regular feedback. Unfortunately, when algorithmic performance management relies purely on employee tracking, it can backfire if AI is seen as Management by Surveillance: The use of AI-driven tools does not necessarily promote positive outcomes when employees interpret these algorithmic tools simply as tools for surveillance.

Edwards *et al.* (2024) found that employee sense-making of an algorithmic system was very important. An AI system intended for performance management was linked to stronger extrinsic motivation when employees interpreted it as evidence of management control but stronger intrinsic motivation and lower emotional exhaustion when employees perceived it as providing useful developmental information. In other words, how employees make sense of technological integration, along with how managers communicate, are critical to employee outcomes, not simply the presence of the technology itself.

2.3 Artificial Intelligence in Education and Training

Artificial intelligence can help us customize workplace learning by recognizing missing skills and training courses. That will become ever more essential now, as digital transformation alters the competencies needed in companies. AI learning can empower employees to discover a need for development, make tailored recommendations, and acquire new skills constantly.

The key role of AI, when it comes to HR, is helping us build human expertise and long-term work readiness among employees. But we should resist thinking employees will instinctively be digital. That, too, must be trained – supplemented with a supportive, experimenting work environment.

2.4 Artificial Intelligence and Employee Engagement

AI may be integrated in order to improve employee engagement, through the use of chatbots, employee feedback systems, sentiment analysis and customized communications. These enable organizations to find re-current employee issues, and furnish HR personnel with workforce information. However, the extraction of data about employee information without transparency has been known to decrease trust. Employees need to have insight on what information is taken, its rationale and the role of the information within the organization's choices.

3. Employee Well-Being in the AI-Enabled Workplace

Employee well-being is understood as a concept with multiple dimensions, referring to employees' experiences (psychological, emotional, social, and professional) in the workplace. In an AI-enhanced workplace, well-being can be affected through manipulation of job demands, job resources, autonomy, workload, meaning, trust, and perceived security. For example, Parker and Grote (2022) have proposed that automation and algorithms may alter job resources such as autonomy, skills utilization, and feedback, as well as job demands such as the monitoring of performance.

Technology use has become crucial for the design of technology-augmented work as well as employee well-being. From their analysis of 31 studies, Rojas *et al.* (2024) suggest that AI is currently being implemented in a number of applications designed to enhance employee well-being: employee mental health monitoring, emotional support, customized well-being plans, assessment and training on psychosocial risks. In other words, the use of AI has shown the potential to support the well-being of employees. But AI could have several undesirable outcomes. Malik *et al.* (2022) find several key themes relating to information safety and security, privacy issues, digital transformation, and job insecurity in AI adopter organizations.

Issues can lead to increased technostress and uncertainty among the workforce. For these reasons, AI has a bifacial potential; it can be a resource, but can also a demand, either because it saves some part of our work, supports in meaningful ways, or, alternatively, because it causes surveillance and increased insecurity and pressure to perform.

4. Algorithmic Management and Employee Autonomy

Algorithmic management is defined as the implementation of algorithms and data-based systems to undertake management tasks commonly performed by human supervisors (Wood & McCarthy, 2020). Parent-Rocheleau and Parker (2022) identified six managerial tasks that could be carried out by an algorithm, including job monitoring, goal setting, performance management, Scheduling, reward management and firing decisions. All of these tasks could indirectly influence both work demands and resources and could potentially affect employee well-being.

Some research has offered interesting accounts of what algorithmic management does, such as Meijerink and Bondarouk (2023), who advanced a “dual face” perspective of algorithmic management, suggesting that this approach could both limit and provide autonomy to workers.

For example, an algorithm may remove worker autonomy by scheduling their work, while on other occasions it may support worker autonomy by providing additional information and resources to them for their task management. Work-place empirical support for autonomy also exists. A cross European study by Kinowska and Sienkiewicz (2023) found that Algorithmic Management could impact employee well-being, and this association with well-being was transmitted through job autonomy and perceived justice. In addition, Zayid *et al.* (2024) analyzed the association between algorithmic management and workforce well-being (burnout and threat perception) based on SDT and COR theory. Therefore, autonomy needs to be investigated as an influential mechanism between AI-based HRM and employee well-being.

5. AI, Fairness, Trust and Employee Experience

5.1 Algorithmic Fairness

Fairness is arguably the biggest concern with the use of AI for human resources management. Many organizations might believe that an algorithm, by operating through math equations, will be more unbiased than an HR human employee. Nevertheless, an algorithm will reflect the historic data upon which it was developed and can therefore replicate the biased decision making of the past. Buiold *et al.* (2024) revealed a list of responsible AI principle with bias, discrimination, justice and trust among those prevalent in the relevant empirical literature.

Qin *et al.*, (2023) provide evidence in their paper of how employees would evaluate the two systems of bias-decision making more closely with employees preferring AI evaluated methods in the contexts they identified. It isn't safe to conclude then that algorithms are either fair or unfair and it would appear they can be anything depending on quality of data, data bias, design and human intervention etc.

5.2 Trust in AI

Trust is an additional factor that drives the acceptability of AI-enabled HR practices by employees. The research done by Xu *et al.* (2024) explored how employees build the initial trust in AI in the HRM domain. It was found through experimental study of 426 participants that

trustworthiness of leader as well as explanations of AI is beneficial to build employee trust in AI and the intention to utilize it. Shamim *et al.* (2023) explored cognitive trust in AI in the developing country context among the front-line employees with the help of interviews and a quantitative survey. They found that trusting employee is necessary for AI adoption. Likewise, Arora and Mittal (2025) studied the 264 HR professionals of Indian E-commerce organizations and concluded that AI-tech trust mediated the relationships of perceptions of changes that are induced by AI in HR function and the intention to adopt AI. These studies are significant for Indian Organizations as it confirms the fact that AI adoption is not only technical decision.

6. Interdisciplinary Perspective

AI-driven HRM requires an interdisciplinary approach because it combines technology with human behaviour and organizational decision-making.

Discipline	Major Concern	Contribution to AI-HRM
Human Resource Management	Workforce practices	Recruitment, performance, learning and retention
Organizational Behaviour	Employee attitudes	Trust, engagement, commitment and resistance
Psychology	Employee experience	Motivation, autonomy, stress and well-being
Information Technology	AI infrastructure	Algorithms, analytics and automation
Business Ethics	Responsible use	Fairness, privacy and accountability
Strategic Management	Organizational capability	Competitive advantage and transformation
Sustainability	Long-term value	Employee well-being and sustainable workforce

Dima *et al.* (2024) show this interdisciplinary very explicitly through their inclusion of literature across disciplines and study of evolutions of roles of employees, managers and HR professionals. Likewise, Pessach *et al.* (2023) use a systematic review to show AI-HRM impacts (at the organizational and employees level) necessitating multidimensional theoretical and managerial lenses.

7. Theoretical Background

7.1. Sociotechnical Systems Viewpoint

Sociotechnical view can be applied as basis for understanding of AI-HRM, as the organization is composed by interrelated socio-techno-technical systems. AI cannot achieve a higher productivity rate unless it takes into account with employee expertise, and, how they fit in company management strategies, structure or organization culture.

AI Technology → HR Practices → Employee Experience → Organizational Outcomes

Outcomes it is argued that, rather than optimizing each element, technology and personnel should go under dual optimization to attain better productivity level.

7.2 Self-Determination Theory

The Self-Determination Theory identifies three fundamental psychological needs for human well-being: autonomy, competence and relatedness. For AI this means, giving valuable information when required and not unnecessarily restricting employees' choices in doing that information. AI help competence when personalizing the information for skills developing purposes. Excessive monitoring could diminish autonomy and interaction replaced by automation could be associated to diminished relatedness. Zayid *et al.* (2024) confirmed the validity of Self Determination theory in relation to algorithmic management and the workforce well-being.

7.3 Capability Approach

Lamers *et al.* (2022) apply the Capability Approach to worker dignity under algorithmic management. Their framework highlights that algorithmic systems can both violate and promote worker dignity depending on whether they expand or restrict workers' capabilities.

This perspective is particularly valuable for AI-HRM because it shifts attention from organizational efficiency alone toward the question of whether employees are actually enabled to develop, participate and exercise meaningful choice at work.

8. Proposed Conceptual Framework

Based on the literature, this chapter proposes a human-centred conceptual framework.

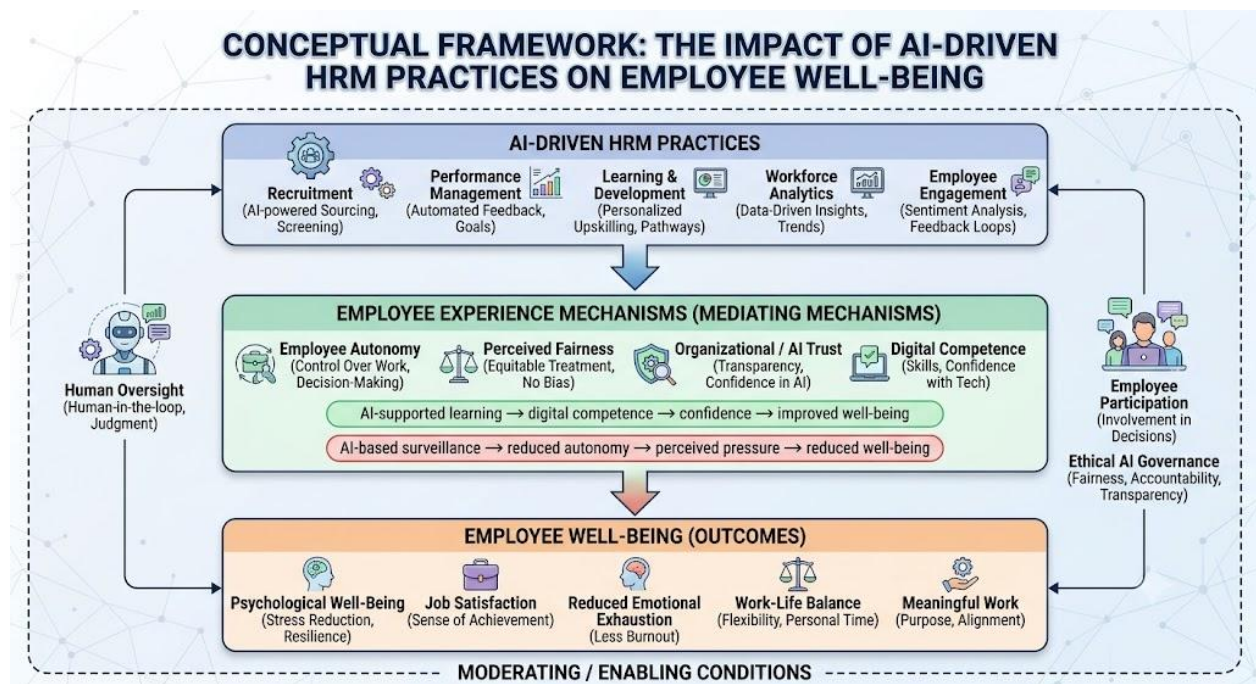


Figure 1: Conceptual Framework

The framework therefore rejects a simple assumption that AI is either beneficial or harmful. Instead, it emphasizes the conditions under which AI produces positive or negative employee outcomes.

9. Research Propositions

The proposed model builds on the following research propositions:

- P1: AI-driven HRM practice significantly affects employee well-being.
- P2: Employee autonomy mediates the impact of AI-driven HRM practice on employee well-being.
- P3: Perceived fairness mediates the relationship between AI-driven HRM practice and employee well-being.
- P4: Organizational trust mediates the relationship between AI-driven HRM practice and employee well-being.
- P5: Digital competence has a positive impact on employee adaptation to AI-driven HRM practice.
- P6: Algorithmic surveillance has a negative effect on the employee's autonomy and employee well-being.
- P7: Human oversight reinforces the positive relationship between responsible AI-HRM practices and employee well-being.
- P8: Employee participation strengthens trust and acceptance toward AI-enabled HRM.

10. Challenges to AI-Driven HRM

Despite the advantages above, several challenges require careful management.

10.1 Algorithmic bias

AI systems often perpetuate and amplify biases inherent in training data, particularly those which relate to recruiting, promotion, and performance management.

10.2 Privacy concerns

HR systems can accumulate vast amounts of employee data which organizations will need strict guidelines for, covering the data which is gathered, stored and used by the HR system.

10.3 Lack of explain ability

Employees and colleagues may become mistrustful when they cannot comprehend how a certain AI system has given particular recommendations or decisions.

10.4 Job insecurity

Employees may also see HR driven by artificial intelligence as a threat when they believe that automation is linked to reducing job numbers or rendering present skills obsolete.

10.5 Over-monitoring of employees

Ongoing observation may produce employees' stress levels and lower job autonomy.

10.6 Skills mismatch

Adoption of AI based technology may require a new set of employee and management technological, analytic, and interpersonal skills for staff.

10.7 Erosion of human relationships

HRM depends on factors such as compassion, communication and relationships, and an excessive use of AI-based approaches will reduce the level of these factors.

11. Managerial Implications

Organization should follow a human-centered AI-HRM strategy rather than a purely automation strategy. Firstly, AI should be leveraged to supplement human decision making, not to take decision making out of the hands of individuals when making critical HR choices.

Secondly, transparency must be established through AI policies that specify the uses of AI, the type of employee data acquired, and the mechanism of how automated decisions will impact on HR actions.

Thirdly, the implementation of AI should involve employee participation in order to facilitate change resistance and build trust, because the workforce are likely to support new technologies, in a context that their purpose is understood by, and employees can discuss concerns on its application. Fourthly, companies must focus on digital reskilling initiatives to enable their workforces understand and embrace new technologies when adopting them. Fifthly, human control must be maintained in key employment-related decision-making processes that affect individuals greatly (i.e. Selection, promotion, disciplinary actions and dismissals).

Lastly, AI implementation should be assessed not based solely on increased productivity and reduced costs but on how employees react through employee-centric factors, e.g.

Self-determination, trust, sense of fairness, job satisfaction, workload, well-being.

12. Future Research Directions

Several opportunities exist for future research.

- Firstly, research should be undertaken more widely in developing economies, especially in India, as context may affect an employee's perception of AI-HRM.
- Secondly, employees may express greater autonomy when engaged with AI-HRM, therefore future research can test if autonomy is a mediator in the link between AI-HRM and employee well-being.
- Thirdly, researchers can explore how organizational trust and perceived fairness function as mediators or moderators between AI-HRM and employee well-being.
- Fourthly, longitudinal research would clarify whether employees' perceptions of AI may evolve over time. Fifthly, sector specific studies are needed to see how different IT, auto, banking, healthcare, education, logistics and manufacturing sector organisations fare under

AI-HRM. Sixthly, it would be useful for researchers to explore differences in AI acceptance, digitally competent and perceive job security among younger vs. older employees.

- Lastly, research should adopt a multi-theoretic approach by incorporating elements of the Self-Determination Theory, the Job Demands-Resources (JDR) Model, sociotechnical system theory, the Resource-Based View (RBV), the Organizational Justice Theory and the Capability Approach.

Conclusion

This study identified that artificial intelligence is changing human resource management by transforming organisation-manager-employee relationship. AI has the potential to enhance efficiency in recruitment, simplify repetitive task, provide personalized development, and facilitate performance management and present organizations with data-based reports. However, to utilize AI in HR organizations are bound to face certain issues related to control, privacy, transparency, trust, security, job loss and employee emotional well-being.

The literature survey suggests that AI affects employee outcomes in different manners which are dual in nature.

AI can act as a work resource when its employees reduced stress from work, easy access to required information or opportunities of development, in contrast; it can become work pressure when employees feel that they are excessively supervised or have limited freedom. For instance, Meijerink & Bondarouk (2023) presented that algorithmic management restrict while it enables employee autonomy at the same time. According to Edwards *et al.* (2024), employee's response to algorithmic HR systems had implications for motivation and emotional exhaustibility. For these reasons, organizations should have not simply assess the efficiency of AI-HRM, they also have to keep on monitor the employee experience by the implementations.

The article advocates the necessity of the development of a framework to ensure effective implementations of AI-HRM, for example; human oversight, visibility and explanation of AI system, equity, employee involvement, Digital literacy & Training and ethical governances.

So future of HRM should not base upon replacing by technology but its technology supporting human by implementing the technology in HR in such way that it augments human judgement, empathy and humanity while enabling the organization maintain responsibility toward workforce. The central hypothesis formulated is: AI provides strategic advantage when it enhances both organizational capabilities and employee well-being.

References

1. Arora, M., & Mittal, A. (2025). Employees' change in perception when artificial intelligence integrates with human resource management: A mediating role of AI-tech trust. *Benchmarking: An International Journal*, 32(6), 1884–1908.
<https://doi.org/10.1108/BIJ-11-2023-0795>

2. Bujold, A., Roberge-Maltais, I., Parent-Rochelleau, X., Boasen, J., Sénécal, S., & Léger, P.-M. (2024). Responsible artificial intelligence in human resources management: A review of the empirical literature. *AI and Ethics*, 4, 1185–1200. <https://doi.org/10.1007/s43681-023-00325-1>
3. Chen, Z. (2022). Collaboration among recruiters and artificial intelligence: Removing human prejudices in employment. *International Journal for Educational and Vocational Guidance*, 25, 135–149. <https://doi.org/10.1007/s10775-022-09704-4>
4. Dima, J., Gilbert, M.-H., Dextras-Gauthier, J., & Giraud, L. (2024). The effects of artificial intelligence on human resource activities and the roles of the human resource triad: Opportunities and challenges. *Frontiers in Psychology*, 15, 1360401. <https://doi.org/10.3389/fpsyg.2024.1360401>
5. Edwards, M. R., Zubielevitch, E., Okimoto, T., Parker, S. K., & Anseel, F. (2024). Managerial control or feedback provision: How perceptions of algorithmic HR systems shape employee motivation, behavior, and well-being. *Human Resource Management*, 63(4), 691–710. <https://doi.org/10.1002/hrm.22218>
6. Ekuma, K. (2024). Artificial intelligence and automation in human resource development: A systematic review. *Human Resource Development Review*, 23(2), 199–229. <https://doi.org/10.1177/15344843231224009>
7. Gong, Q., Fan, D., & Bartram, T. (2025). Integrating artificial intelligence and human resource management: A review and future research agenda. *The International Journal of Human Resource Management*, 36(1), 103–141. <https://doi.org/10.1080/09585192.2024.2440065>
8. Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at work: The new contested terrain of control. *Academy of Management Annals*, 14(1), 366–410. <https://doi.org/10.5465/annals.2018.0174>
9. Kinowska, H., & Sienkiewicz, Ł. J. (2023). Influence of algorithmic management practices on workplace well-being: Evidence from European organisations. *Information Technology & People*, 36(8), 21–42. <https://doi.org/10.1108/ITP-02-2022-0079>
10. Lamers, L., Meijerink, J., Jansen, G., & Boon, M. (2022). A capability approach to worker dignity under algorithmic management. *Ethics and Information Technology*, 24, Article 10. <https://doi.org/10.1007/s10676-022-09637-y>
11. Lu, Y., Yang, M. M., Zhu, J., & Wang, Y. (2024). Dark side of algorithmic management on platform worker behaviors: A mixed-method study. *Human Resource Management*, 63(3), 477–498. <https://doi.org/10.1002/hrm.22211>

12. Malik, N., Tripathi, S. N., Kar, A. K., & Gupta, S. (2022). Impact of artificial intelligence on employees working in Industry 4.0-led organizations. *International Journal of Manpower*, 43(2), 334–354. <https://doi.org/10.1108/IJM-03-2021-0173>
13. Meijerink, J., & Bondarouk, T. (2023). The duality of algorithmic management: Toward a research agenda on HRM algorithms, autonomy and value creation. *Human Resource Management Review*, 33(1), 100876. <https://doi.org/10.1016/j.hrmr.2021.100876>
14. Möhlmann, M., Alves de Lima Salge, C., & Marabelli, M. (2023). Algorithm sensemaking: How platform workers make sense of algorithmic management. *Journal of the Association for Information Systems*, 24(1), 35–64. <https://doi.org/10.17705/1jais.00774>
15. Parker, S. K., & Grote, G. (2022). Automation, algorithms, and beyond: Why work design matters more than ever in a digital world. *Applied Psychology*, 71(4), 1171–1204. <https://doi.org/10.1111/apps.12241>
16. Parent-Rochelleau, X., & Parker, S. K. (2022). Algorithms as work designers: How algorithmic management influences the design of jobs. *Human Resource Management Review*, 32(3), 100838. <https://doi.org/10.1016/j.hrmr.2021.100838>
17. Pereira, V., Hadjielias, E., Christofi, M., & Vrontis, D. (2021). A systematic literature review on the impact of artificial intelligence on workplace outcomes: A multi-process perspective. *Human Resource Management Review*, 33(1), 100857. <https://doi.org/10.1016/j.hrmr.2021.100857>
18. Pessach, D., Shmueli, E., Singer, G., Ben-Gal, I., & others. (2023). Artificial intelligence (AI)-assisted HRM: Towards an extended strategic framework. *Human Resource Management Review*, 33(1), 100940. <https://doi.org/10.1016/j.hrmr.2022.100940>
19. Qin, S., Jia, N., Luo, X., Liao, C., & Huang, Z. (2023). Perceived fairness of human managers compared with artificial intelligence in employee performance evaluation. *Journal of Management Information Systems*, 40(4), 1039–1070. <https://doi.org/10.1080/07421222.2023.2267316>
20. Rojas, R., et al. (2024). The role of artificial intelligence in improving workplace well-being: A systematic review. *Businesses*, 4(3), 389–410. <https://doi.org/10.3390/businesses4030024>
21. Shamim, S., Yang, Y., Zia, N. U., Khan, Z., & Shariq, S. M. (2023). Mechanisms of cognitive trust development in artificial intelligence among frontline employees: An empirical examination from a developing economy. *Journal of Business Research*, 162, 113–? [Article 114168]. <https://doi.org/10.1016/j.jbusres.2023.114168>
22. van Zoonen, W., Sivunen, A. E., & Treem, J. W. (2024). Algorithmic management of crowdworkers: Implications for workers' identity, belonging, and meaningfulness of work. *Computers in Human Behavior*, 152, 108089. <https://doi.org/10.1016/j.chb.2023.108089>

23. Xu, Y., Huang, Y., Wang, J., & Zhou, D. (2024). How do employees form initial trust in artificial intelligence: Hard to explain but leaders help. *Asia Pacific Journal of Human Resources*, 62(3), e12402. <https://doi.org/10.1111/1744-7941.12402>
24. Zayid, H., Alzubi, A., Berberoğlu, A., & Khadem, A. (2024). How do algorithmic management practices affect workforce well-being? A parallel moderated mediation model. *Behavioral Sciences*, 14(12), 1123. <https://doi.org/10.3390/bs14121123>
25. Zhang, A., Boltz, A., Wang, C. W., & Lee, M. K. (2022). Algorithmic management reimagined for workers and by workers: Centering worker well-being in gig work. *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, 1–20. <https://doi.org/10.1145/3491102.3501866>

IOT-BASED GREENHOUSE CONTROLLING AND MONITORING SYSTEM USING ARDUINO UNO AND ESP8266

Shalu C., Mohammad Altamas and Debika Chaudhuri

Department of Electronics and Communication Engineering,
School of Engineering & Technology, IFTM University, Moradabad, U.P., India

*Corresponding author E-mail: shalu268@gmail.com

Abstract

Greenhouse cultivation depends on maintaining a narrow and crop-appropriate range of environmental conditions. In small and medium-scale farming environments, continuous manual supervision can become difficult because temperature, humidity, soil moisture, and illumination vary during the day and may change rapidly under external weather disturbances. This paper presents the design, implementation, and functional evaluation of a low-cost Internet of Things (IoT)-enabled greenhouse monitoring and control prototype developed from the underlying M.Tech project. The system combines an Arduino UNO controller, ESP8266 NodeMCU wireless module, DHT11 temperature-humidity sensor, soil-moisture sensor, light-dependent resistor (LDR), relay interface, DC motor-based actuator arrangement, LCD display, and Blynk mobile dashboard. Sensor observations are continuously acquired and compared with predefined control thresholds. When a parameter moves outside the desired operating condition, the controller triggers the corresponding actuator, while the ESP8266 transfers measurements to the remote monitoring platform. The prototype was repeatedly tested under normal and deliberately varied environmental conditions. The evaluation confirmed functional sensing, threshold-based actuation, local display, remote visualization, stable Wi-Fi communication during the test sessions, and reduced dependence on manual intervention. In the laboratory demonstration, an increase in temperature activated the ventilation surrogate, low soil moisture initiated the irrigation surrogate, and excessive humidity produced a ventilation response. Variations in ambient light were successfully detected and visualized. The work demonstrates that a practical greenhouse automation platform can be assembled with widely available and relatively inexpensive hardware while preserving a modular architecture that can later support additional sensors, predictive analytics, alternative communication technologies, and renewable-energy operation. The study also identifies limitations associated with low-cost sensing accuracy and dependence on Internet connectivity and therefore positions the presented prototype as a validated proof-of-concept rather than a fully commercial greenhouse controller.

Keywords: Internet of Things, Smart Greenhouse, Precision Agriculture, Arduino UNO, ESP8266, NodeMCU, DHT11, Soil Moisture, LDR, Blynk, Automated Irrigation, Environmental Monitoring.

1. Introduction

Agricultural production is increasingly being influenced by the availability of sensing, communication, and embedded control technologies. Conventional cultivation practices depend heavily on periodic human inspection. Although this approach can be adequate for small plots, it becomes less effective when crops are grown inside protected environments where climatic variables need to be regulated repeatedly and at short time intervals. A greenhouse is intended to create a controlled growing environment, but the benefit of protection is realized only when temperature, relative humidity, soil water status, and light are kept within acceptable ranges. The thesis underlying this paper identifies these parameters as the central environmental variables for the proposed greenhouse controller and further emphasizes that remote access can reduce the need for constant physical presence at the site.

The development of low-cost microcontrollers and wireless communication modules has made it possible to build greenhouse controllers that are considerably more accessible than industrial automation platforms. The reviewed literature in the thesis includes earlier systems based on ZigBee, GSM, Android applications, Raspberry Pi, and Arduino-oriented sensing platforms. These studies show a progression from wired or workstation-centered monitoring toward distributed and remotely accessible greenhouse management. The present work follows that direction but concentrates on a compact architecture that combines local sensing and control with Internet-connected supervision through the ESP8266 NodeMCU and Blynk platform. [1-14]

The central research problem addressed in this study is therefore not simply how to measure greenhouse variables, but how to connect sensing, decision-making, actuation, and remote visualization into one low-cost operational loop. In the implemented prototype, the Arduino UNO handles the sensor and control logic, the ESP8266 provides wireless connectivity, relay circuitry isolates the low-power control signals from actuator loads, and the Blynk application presents the measured variables to the user. This arrangement creates a feedback loop in which a physical condition is sensed, interpreted against a predefined threshold, translated into an actuator command, and simultaneously exposed to a remote user. [15-19]

1.1 Research Problem

Many greenhouse systems reported in the literature concentrate on a subset of environmental parameters, require specialized communication infrastructure, or rely on a computer-based supervisory station. Earlier ZigBee systems, for example, demonstrated distributed sensor networking and remote monitoring, but they often included gateway computers or dedicated wireless network infrastructure. Other systems emphasized Android monitoring, SMS alerts, or specific control functions. The thesis project was motivated by the practical need for a simpler arrangement suitable for educational, laboratory, and small-scale agricultural use: a controller that can sense several greenhouse parameters, automatically react to abnormal conditions, and

make the state of the greenhouse visible on a smartphone without a permanently attached desktop computer. [1-8]

1.2 Research Objectives

The objectives of the study are derived directly from the implemented prototype and its documented testing:

- To monitor temperature and relative humidity continuously using a DHT11 sensor;
- To detect soil moisture conditions and trigger automated irrigation logic when moisture falls below the programmed threshold;
- To measure changes in ambient light using an LDR and provide a basis for shading or lighting control;
- To connect local monitoring and control logic to the Internet through an ESP8266 nodemcu module;
- To present live sensor information locally on an LCD and remotely through the Blynk mobile application;
- To evaluate the functional response of the integrated prototype under normal and deliberately altered environmental conditions; and
- To examine the practicality, scalability, and limitations of a low-cost greenhouse automation approach.

1.3 Contributions of the Study

The principal contribution of this work is an integrated proof-of-concept rather than a new sensing device or a new control algorithm. Its value lies in the practical coupling of four environmental observations with automatic responses and cloud-connected visualization in one compact prototype. The design also demonstrates a clear migration path toward more advanced smart-farming systems. According to the thesis, additional CO₂, pH, nutrient, rainfall, and weather-related sensing can be added without redesigning the complete architecture, while future predictive functions can be introduced through machine learning.

2. Related Work and Research Gap

The literature summarized in the thesis shows that greenhouse automation has been studied through several technology generations. Jian Song proposed a ZigBee wireless measurement and control architecture using a central controller and distributed sensor nodes for greenhouse monitoring. The system was tested with multiple nodes and reported measurement errors within 3% for temperature, humidity, and illuminance. This work demonstrates the feasibility of wireless environmental sensing at greenhouse scale, but it also illustrates the hardware and network complexity of an explicitly designed wireless sensor network. [1]

Ai and Chen focused on the use of a mobile phone as a monitoring terminal, illustrating the importance of mobile access in greenhouse management. Related Android-oriented work

extended monitoring and control to temperature, soil moisture, and humidity, while other systems combined sensors, Internet connectivity, and smartphone interfaces. These approaches establish that remote accessibility is useful because the operator does not have to remain physically present in the greenhouse. [2,4,8,10]

Wireless sensing approaches based on ZigBee and sensor networks have also emphasized the rapid collection of environmental measurements and centralized analysis. Several studies in the thesis describe architectures with dedicated wireless nodes, gateways, PCs, or laboratory software such as LabVIEW. Such designs can be appropriate for large installations, but a smaller agricultural or educational deployment may benefit from an architecture with fewer intermediary components and direct Wi-Fi connectivity. [1,3,15,18,19]

Arduino-centered research has explored temperature, humidity, soil moisture, illumination, and in some cases soil pH. Vimal and Shivaprakasha, for example, described an Arduino-based greenhouse environment monitoring and control platform that used multiple environmental sensors and remote communication. Other work has examined automated greenhouse monitoring, intelligent greenhouse control, and soil or crop-related parameter analysis. These studies support the feasibility of the sensor-control paradigm adopted here. [7,8,11,12]

A second research direction concerns climate control in tropical environments. The thesis literature includes work in which real-time and historical weather data are used to support control algorithms for greenhouse temperature and humidity. Such approaches point toward a transition from simple threshold control to adaptive and predictive control. The current prototype remains at the reactive stage: it compares sensor values with predefined thresholds and triggers an actuator when a limit is crossed. That makes the system transparent and easy to implement, but it also creates a clear opportunity for future research involving prediction, crop-specific optimization, and learning-based control. [9,17]

Taken together, the reviewed studies reveal a practical gap for compact low-cost systems that provide four-way environmental sensing, local automatic control, and Internet-based mobile visualization within a single demonstrator. The present work addresses this gap through the integration of Arduino UNO, NodeMCU/ESP8266, DHT11, soil-moisture and LDR sensing, relay actuation, LCD output, and Blynk monitoring. The contribution is therefore best understood as an integrated and reproducible system architecture with functional experimental validation rather than as a claim of algorithmic novelty.

3. Proposed System Architecture

The proposed greenhouse controller is organized as four interacting layers: sensing, local decision and control, actuation, and remote monitoring. The sensing layer acquires temperature, humidity, soil moisture, and light information. The local control layer interprets these measurements on the microcontroller side. The actuation layer consists of relay-controlled loads

represented in the prototype by DC motors. Finally, the communication layer uses the ESP8266 NodeMCU to transport measurements to the Blynk IoT application. The thesis also includes an LCD for direct local observation of the measured values.

Figure 1: Functional architecture of the proposed IoT greenhouse controller

Environmental conditions	Sensing	Controller	Actuation	User / Cloud
Temperature Humidity Soil moisture Light	DHT11 Soil moisture sensor LDR	Arduino UNO Threshold logic	Relay module DC motor / pump / fan surrogate	ESP8266 NodeMCU Blynk dashboard
Physical greenhouse	Measurements	Decision	Control action	Remote visibility
Heat / dryness / high humidity / high light	Analog + digital readings	Compare with preset limits	Activate or deactivate load	Live mobile monitoring

3.1 Hardware Components

The hardware was selected around availability, ease of interfacing, and suitability for a prototype. The Arduino UNO serves as the main microcontroller and provides the local processing platform for sensor inputs and actuator control. The board is based on the ATmega328P family and, as described in the thesis, provides digital and analog I/O suitable for connecting the required sensors and peripherals.

The ESP8266 NodeMCU acts as the wireless IoT interface. The thesis describes it as a low-cost platform built around the ESP8266 with Wi-Fi connectivity, GPIO, ADC, PWM, and serial communication capabilities. These functions make it suitable for moving the greenhouse prototype from a purely local controller to an Internet-connected system. The selected architecture uses Wi-Fi connectivity for regular transmission of sensor information and remote supervision.

Temperature and humidity are captured with the DHT11. The thesis reports an operating temperature range of 0-50 °C and a relative-humidity range of 20-95%, with the stated sensor accuracies of approximately ± 2 °C for temperature and $\pm 5\%$ for humidity. The thesis also notes its low sampling rate, which is one reason the device is appropriate for a low-cost prototype but not necessarily ideal for demanding industrial control.

The soil-moisture sensor provides the variable required for automated irrigation. In the prototype, the sensor output is interpreted against a threshold so that irrigation can be initiated when the soil is considered too dry. The LDR provides a simple indication of ambient light intensity and is used as the input for the light-control decision. Relay modules provide electrical

switching between the low-power control electronics and the actuator side. The prototype uses DC motors to simulate the physical actions that would otherwise be performed by a fan, shade mechanism, or water-pumping assembly.

Table 2: Principal hardware and functional roles in the prototype

Component	Function	Role In Prototype	Design Consideration
Arduino UNO	Local embedded control	Reads sensors and executes control logic	Low cost, simple programming, adequate I/O
ESP8266 NodeMCU	Wi-Fi / IoT communication	Transfers sensor status to Blynk	Requires reliable Wi-Fi for remote operation
DHT11	Temperature + humidity sensing	Climate feedback	Low-cost but limited precision and sampling rate
Soil-moisture sensor	Soil water-status indication	Irrigation trigger	Threshold calibration is environment dependent
LDR	Ambient light sensing	Light/shade decision input	Best suited to relative light changes
Relay module	Electrical switching	Interfaces control logic with loads	Provides switching isolation
DC motor	Actuator surrogate	Represents fan/pump/shade action	Used for laboratory demonstration
LCD + Blynk	Local + remote display	Shows live parameter status	Improves operator visibility

3.2 Software and Communication Layer

The software environment described by the thesis consists of the Arduino IDE and embedded C-style programming for the controller. The control strategy is event-oriented and threshold-based. Rather than estimating a future climate state, the controller reacts to present measurements: when a monitored variable crosses a predefined operating limit, the corresponding actuator is commanded. The ESP8266 provides the network connection required for the remote dashboard. The thesis reports that the Blynk interface displays the readings continuously and offers graphical visualization of changes over time.

3.3 Control Logic

The control mechanism can be summarized as a sequence of measurement, comparison, action, and verification. At each sampling cycle, the system acquires the four environmental variables. The controller compares the current values with the programmed thresholds. If temperature rises above the intended range, the ventilation actuator is enabled. If humidity becomes excessive, the ventilation response is also activated. If soil moisture falls below the acceptable level, the irrigation actuator is enabled until the desired moisture condition is restored. Changes in light

intensity are mapped to a possible shading or artificial-light response. The thesis states that the DC motor is used as a laboratory substitute for the actual greenhouse fan, shade, or water-outlet mechanism.

Table 3: Threshold-oriented control logic and laboratory observations

Monitored Variable	Trigger Condition Described in Thesis	Prototype Response	Observed Evaluation
Temperature	Temperature crosses programmed threshold	Relay activates DC motor as fan/exhaust surrogate	Motor activated when heat was intentionally increased
Humidity	Humidity exceeds programmed limit	Ventilation mechanism / motor response	Humidity variation produced control response
Soil moisture	Moisture falls below programmed threshold	Watering mechanism / motor response	Dry-soil condition triggered automatic watering action
Light intensity	Light exceeds programmed threshold or varies significantly	Basis for shading / artificial-light response	LDR changes were detected and visualized
Connectivity	Wi-Fi connection available	ESP8266 uploads readings to Blynk	Stable communication reported during testing

4. Materials and Methods

The research follows a prototype-development and functional-testing methodology. The aim was to construct an operational greenhouse monitoring-and-control loop and then verify whether the individual sensing channels, communication path, and actuator responses behaved as intended. Because the thesis does not provide a statistically structured dataset or a long-duration agronomic field trial, the evaluation is deliberately interpreted as a laboratory proof-of-concept. The reported claims therefore concern functional behavior, remote accessibility, and control operation rather than quantified crop-yield improvement.

4.1 Prototype Construction

The prototype was assembled by connecting the environmental sensors to the embedded controller and interfacing the actuator loads through relay modules. The DHT11 provides temperature and humidity information; the soil-moisture sensor provides the irrigation-related input; and the LDR provides the light-related input. The LCD serves as a local human-readable display. The ESP8266 NodeMCU establishes the wireless link required for the Blynk dashboard. The final unit therefore combines local embedded control with Internet-connected monitoring while retaining a simple physical wiring structure.

4.2 Experimental Procedure

Testing was performed repeatedly under normal environmental conditions and under deliberately altered conditions. This strategy is particularly appropriate for a threshold-driven controller because the critical test is not merely whether a sensor displays a value, but whether a meaningful environmental disturbance propagates through the entire sensing-decision-actuation chain. For temperature testing, the thesis reports that an external heat source was used to raise the surrounding temperature. When the programmed threshold was crossed, the controller activated the DC motor through the relay and subsequently turned it off after the temperature returned to the acceptable range.

Humidity was evaluated by observing changes in the measured relative humidity and verifying that the control logic produced the intended ventilation response once the predefined limit was exceeded. Soil moisture was tested by allowing the soil condition to fall below the programmed threshold. The controller then activated the irrigation surrogate and stopped it after the desired moisture state was restored. Light behavior was tested through changes in ambient illumination, with the LDR readings displayed on the mobile dashboard.

The communication path was evaluated simultaneously with the physical control tests. The LCD provided an immediate local indication of sensor readings, while the Blynk application displayed the same measurements remotely. This dual-observation approach helped verify that the embedded measurement stream was reaching both the local display and the remote interface. The thesis reports stable Wi-Fi connectivity during the test sessions and no noticeable delay in the transfer of sensor data to the remote platform.

4.3 Evaluation Criteria

The system was evaluated using five functional criteria: (1) sensor observability, meaning that relevant environmental changes were detected; (2) threshold response, meaning that abnormal states generated an actuator command; (3) recovery behavior, meaning that the actuator could be switched off after the monitored variable returned to an acceptable region; (4) remote accessibility, meaning that readings were visible through the Blynk dashboard; and (5) operational practicality, meaning that the prototype reduced the need for repeated manual inspection. These criteria reflect the actual evidence available in the thesis and avoid introducing performance metrics that were not experimentally recorded.

5. Results

5.1 Integrated System Operation

The integrated prototype was reported to operate successfully under the test conditions. Sensor readings were displayed on both the LCD and the Blynk application, and the ESP8266 enabled real-time transmission of the measurements over Wi-Fi. The thesis describes the resulting system as capable of monitoring environmental conditions remotely from any location with Internet

connectivity. This is a meaningful systems-level outcome because greenhouse automation becomes more useful when the sensing and control process can continue without requiring the operator to remain beside the physical unit.

5.2 Temperature Regulation Response

Temperature control produced one of the clearest end-to-end demonstrations. Under normal conditions, the DHT11 measurements remained within the desired operating range. The test temperature was then deliberately increased using an external heat source. Once the measurement crossed the programmed threshold, the NodeMCU/controller chain activated the motor through the relay. The motor represented the ventilation or exhaust-fan action. After the temperature returned toward the acceptable range, the controller switched the motor off. The experiment therefore verified not only threshold detection but also the complete action-and-recovery cycle.

From a control-engineering perspective, this result confirms that the prototype behaves as a reactive feedback system. The control decision is generated from the current sensed state rather than by human interpretation. The architecture is simple, which is an advantage for transparency and implementation, but the thesis does not report a numerical settling time, temperature error, overshoot, or controller stability metric. Consequently, the result should be stated as successful functional regulation rather than as a quantified temperature-control performance benchmark.

5.3 Humidity Regulation

The humidity channel was evaluated under changing atmospheric conditions. When humidity exceeded the programmed limit, the control logic-initiated ventilation. The thesis reports that the system-maintained humidity within the desired range during the test scenario and notes the practical importance of avoiding excessive moisture because high humidity can create conditions favorable to fungal growth and other crop problems. The key result is therefore the successful integration of humidity sensing into the same automatic control loop used for temperature.

5.4 Soil-Moisture-Based Irrigation

The soil-moisture channel produced the irrigation response. As the soil condition became dry and the sensor reading moved below the programmed threshold, the controller energized the relay and activated the watering surrogate. The action continued until the desired soil condition was restored, at which point the motor was turned off. In operational terms, the mechanism changes irrigation from a purely time-based or manually triggered activity into a condition-based activity. This is important because the system supplies water when the observed soil state demands it rather than simply following a fixed schedule.

The thesis further interprets this behavior as a means of reducing unnecessary water use and minimizing manual effort. That interpretation is reasonable at the prototype level because the pump response is explicitly linked to the soil condition. However, the paper does not claim a

measured percentage reduction in water consumption because the thesis does not present a controlled water-use baseline or volumetric irrigation dataset. The result is consequently described as potential for resource-efficient operation, supported by the observed on-demand control behavior.

5.5 Light Monitoring

The LDR channel successfully responded to changes in ambient illumination. During the testing process, higher light intensity generated corresponding sensor readings that were transferred to the Blynk dashboard. The architecture was designed so that the light threshold could be used to activate a shading mechanism or an artificial-lighting response depending on the crop requirement. In the laboratory prototype, the motor can therefore be interpreted as a demonstration actuator rather than as a final commercial shading assembly.

5.6 Remote Monitoring and Communication

The Blynk application provided the remote observation layer of the system. According to the thesis, temperature, humidity, soil moisture, and light readings were updated continuously on the smartphone dashboard, with graphical visualization of environmental changes. Stable Wi-Fi connectivity was observed during the test sessions, allowing the NodeMCU to communicate with the cloud platform without interruption. The reported response was satisfactory for the laboratory setting.

Remote monitoring changes the operational model of the greenhouse. Instead of requiring the farmer or operator to walk to the greenhouse to inspect every variable, the state of the environment can be checked through a mobile interface. The control loop itself remains automatic, so the smartphone serves primarily as an observation and management interface rather than as the sole decision-maker. This distinction is valuable because loss of smartphone access does not necessarily stop local threshold-based actuation, although loss of connectivity does affect remote visibility.

5.7 Reduction in Manual Intervention

The thesis reports that the prototype reduces manual supervision because routine inspection and actuation are performed automatically. In a conventional workflow, a user would need to check environmental conditions, decide whether irrigation or ventilation is required, and then operate the corresponding equipment. In the proposed system, the sensing and decision logic performs these steps continuously. This reduces repetitive human effort and can also reduce delays caused by intermittent manual observation.

5.8 Resource Utilization

An additional qualitative result concerns resource use. The thesis reports that water was supplied only when the soil-moisture condition indicated insufficient moisture and that electrical loads were activated when required by the threshold logic. The architecture therefore encourages

demand-based operation rather than continuous operation. It is important, however, to distinguish between a mechanism that is designed to avoid unnecessary resource consumption and a statistically measured saving. The current study demonstrates the former; a future field study would be required to quantify the latter.

Table 4: Functional test scenarios and observed behavior reported in the thesis

Test Scenario	Injected / Observed Condition	Expected Action	Documented Result
Normal operation	Environmental variables within desired range	Continue sensing; keep actuators inactive unless required	System operated and displayed readings
High temperature	Temperature intentionally increased	Activate ventilation surrogate	Motor activated; later switched off after recovery
High humidity	Humidity raised above programmed limit	Activate ventilation response	Humidity control response observed
Dry soil	Moisture falls below programmed threshold	Activate irrigation surrogate	Watering action initiated and stopped after recovery
High / changing light	LDR detects increased illumination	Enable shading/lighting logic as configured	Light changes detected and visualized
Remote observation	Wi-Fi available	Transmit readings to Blynk	Continuous mobile updates and graphical display observed

6. Discussion

The results indicate that the prototype successfully closes the fundamental loop required for basic greenhouse automation: measurement is translated into a control decision, the decision drives an actuator, and the system state is exposed to both local and remote users. This integration is more important than any individual component because an isolated temperature sensor or Wi-Fi module alone does not constitute greenhouse automation. The architecture demonstrates that low-cost embedded hardware can support a multi-parameter feedback system without requiring a dedicated industrial controller or a permanent supervisory computer.

The most important practical strength is modularity. Each sensor provides one environmental input and each actuator represents one control pathway. Because the system is not built around a single monolithic sensing component, additional variables can be added later. The thesis explicitly identifies CO₂, pH, rainfall, nutrient, and weather-related sensing as possible extensions. This modular approach is consistent with the direction of smart agriculture, where a

greenhouse can evolve from basic reactive automation toward more comprehensive environment-and-crop intelligence.

A second strength is the combination of local and remote operation. The LCD provides immediate visibility at the device, while Blynk provides remote access through the smartphone. This is useful in situations where the network is available but the operator is physically distant. It also provides a basic form of redundancy at the interface level: local operation can still be observed even when cloud visibility is unavailable, while remote monitoring provides convenience when the site cannot be inspected directly.

The prototype also demonstrates a reasonable pathway from reactive to predictive control. At present, a threshold is the decision boundary. A more advanced controller could learn the relationship between environmental variables, external weather, time of day, and crop requirements. The thesis identifies machine learning as a future direction for predicting unfavorable conditions and improving resource use. Such an extension could eventually enable anticipatory control, for example by increasing ventilation before a forecast temperature peak rather than waiting until the temperature has already crossed a limit.

6.1 Comparison with Earlier Approaches

Compared with early ZigBee greenhouse systems, the present architecture reduces the dependence on a dedicated wireless sensor network by using ESP8266 Wi-Fi connectivity for the prototype. Compared with GSM-based approaches, it emphasizes continuous Internet-connected visualization rather than SMS-only alerts. Compared with workstation-centered systems using a host PC or LabVIEW, the proposed arrangement is more compact and smartphone-oriented. These differences do not make the current system universally superior; instead, they reflect a deliberate choice for low-cost prototyping and simpler deployment. The literature itself shows that different communication technologies remain appropriate for different greenhouse scales and deployment environments. [1,4,8,15,18,19]

The main difference is the balance between simplicity and sophistication. The proposed system deliberately uses a transparent threshold controller, which makes debugging and demonstration straightforward. More sophisticated studies in the literature use fuzzy control, historical data, prediction, or distributed sensor networks. Those methods may provide stronger adaptability, but they also increase implementation and maintenance complexity. The current work therefore occupies a practical middle ground: more capable than a manual greenhouse, but simpler than a large predictive or industrial control platform.

6.2 Research Significance

From a research and engineering perspective, the significance of the prototype lies in demonstrating a coherent path from sensing to action. The individual technologies used are established, but their integration into an affordable greenhouse controller creates a useful

reference implementation for small-scale precision agriculture. The system can be reproduced, modified, and extended because the hardware components and control relationships are explicit. The thesis further demonstrates that the same architecture can be adapted to other environmental monitoring tasks where conditions are measured continuously and actions are triggered by thresholds.

7. Limitations

The study has several limitations that should be stated clearly. First, the prototype relies on low-cost sensors. The thesis specifically notes that DHT11 accuracy may vary under extreme conditions. This means that the measured values should not be interpreted as equivalent to readings from industrial-grade environmental instrumentation. The low sampling frequency and sensor resolution also limit the precision with which fast environmental changes can be characterized.

Second, the reported evaluation is primarily functional and qualitative. Although the system was tested repeatedly under altered environmental conditions, the thesis does not provide a large logged dataset, statistical error analysis, numerical response-time distribution, power-consumption measurements, or a controlled crop-yield comparison. Therefore, the study establishes feasibility and functionality but does not yet establish statistically significant agronomic benefit. A journal extension should maintain this distinction to avoid overstating the results.

Third, remote monitoring depends on Internet connectivity. The thesis reports stable Wi-Fi communication in the laboratory tests but recognizes that rural connectivity limitations may produce communication delays or loss of remote access. In an agricultural deployment, this is an important engineering constraint. A controller that relies too heavily on the cloud may become less dependable when connectivity is intermittent. Local fail-safe control and buffered data storage would therefore be useful additions in future implementations.

Fourth, the actuators in the prototype are represented by DC motors for demonstration purposes. A field deployment would require properly rated fans, pumps, valves, shade mechanisms, electrical protection, and appropriate power electronics. The behavior of the laboratory prototype should therefore be interpreted as proof of the control logic rather than as a complete production-ready actuator design.

8. Future Research Directions

The next research stage should move beyond simple threshold control toward crop-aware adaptive management. The thesis proposes integrating artificial intelligence, machine learning, image processing, additional environmental sensors, wireless sensor networks, cloud computing, renewable energy, and stronger cybersecurity. These directions can be organized into four broad

themes: richer sensing, predictive intelligence, scalable communication, and energy-aware deployment.

Richer sensing would include CO₂ concentration, soil pH, nutrient indicators, rainfall, and crop-specific variables. These measurements could allow a controller to distinguish between conditions that are merely acceptable and conditions that are optimal for a specific crop. Camera-based computer vision could extend the system further by detecting visible disease symptoms, pest damage, plant growth patterns, or nutrient stress. Such data could complement the environmental sensors and create a fuller representation of crop health.

Predictive intelligence is another major opportunity. Instead of waiting for temperature or humidity to cross a threshold, the controller could use historical sensor data together with weather information to estimate future conditions. Machine-learning models could then optimize irrigation, ventilation, and lighting while accounting for time-of-day patterns and crop requirements. This would turn the current reactive system into a predictive decision-support platform. The thesis explicitly identifies this transition as a future possibility.

For larger installations, the communication layer could be redesigned around ZigBee or LoRaWAN sensor networks. A multi-node architecture would allow the greenhouse to be divided into zones so that environmental variation within a large structure could be captured more accurately. Local edge processing could reduce the amount of data transmitted to the cloud and would also preserve essential control behavior during temporary Internet outages.

Energy autonomy is also relevant for rural agricultural deployments. Solar power, low-power sensing schedules, sleep modes, and efficient communication could reduce dependence on grid power. Finally, cybersecurity should be treated as a first-class engineering requirement as the system grows. Authentication, secure data transfer, device management, and controlled actuator access become increasingly important when greenhouse equipment is connected to the Internet.

Conclusion

This paper presented an expanded research-oriented account of an IoT-based greenhouse monitoring and control prototype developed using Arduino UNO, ESP8266 NodeMCU, DHT11 temperature and humidity sensing, soil-moisture sensing, LDR-based light monitoring, relay switching, a DC motor actuator arrangement, LCD output, and the Blynk mobile platform. The system was designed to reduce dependence on continuous manual supervision by connecting environmental sensing with automatic threshold-based responses and remote observation. The documented laboratory tests confirmed that the prototype can detect meaningful changes in temperature, humidity, soil moisture, and light, trigger the corresponding actuator behavior, and communicate sensor information to a smartphone dashboard.

The work demonstrates the practical value of low-cost embedded IoT technologies in greenhouse automation, particularly for educational, experimental, and small-scale agricultural settings. Its

most important strength is the integration of sensing, decision-making, actuation, and remote monitoring into a single modular architecture. At the same time, the study recognizes that low-cost sensing accuracy, network dependence, and the absence of long-term quantitative agronomic measurements limit the strength of the current claims. The prototype should therefore be viewed as a validated foundation for further investigation rather than a final commercial solution. Future work that adds calibrated sensors, larger datasets, predictive control, field experiments, scalable wireless networking, cybersecurity, and renewable energy could transform the prototype into a more accurate and resilient smart-farming platform.

References

1. Song, J. (2010). Greenhouse monitoring and control system based on Zigbee wireless sensor network. In *International Conference on Electrical and Control Engineering* (pp. 2785–2788). IEEE Computer Society.
2. Piyare, R. (2013). Internet of Things: Ubiquitous home control and monitoring system using Android based smart phone. *International Journal of Internet of Things*, 2(1), 5–11.
3. Li, G., Zhang, W., & Zhang, Y. (2014). A design of the IoT gateway for agricultural greenhouse. *Sensors & Transducers*, 172(6).
4. Dan, L., Xin, C., Chongwei, H., & Liangliang, J. (2015). Intelligent agriculture greenhouse environment monitoring system based on IoT technology. In *2015 International Conference on Intelligent Transportation, Big Data and Smart City* (pp. 487–490).
5. Balakrishna, K., Nethravathi, S. N., & Krishna, H. (2016). Real-time soil monitoring system for the application of agriculture.
6. Liu, D., Sun, J., Yu, Y., & Xiang, J. (2016). Precise agricultural greenhouses based on the IRT and fuzzy control. In *International Conference on Intelligent Transportation, Big Data & Smart City*.
7. Vishwakarma, A., Sahu, A., Sheikh, N., Payasi, P., Rajput, S. K., & Srivastava, L. (2020). IoT based greenhouse monitoring and controlling system. In *2020 IEEE Students Conference on Engineering & Systems (SCES)* (pp. 1–6). IEEE. <https://doi.org/10.1109/SCES50439.2020.9236693>
8. Danita, M., Mathew, B., Shereen, N., Sharon, N., & Paul, J. J. (2018). IoT based automated greenhouse monitoring system. In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1933–1937). IEEE. <https://doi.org/10.1109/ICCONS.2018.8662911>
9. Ullah, W., Mortuza, M., Kabir, M., Ahmed, Z., Supta, S., Das, P., & Hossain, S. (2019). Internet of Things based smart greenhouse: Remote monitoring and automatic control. *DEStech Transactions on Environment, Energy and Earth Sciences*. <https://doi.org/10.12783/dteees/iceee2018/27803>

10. Satpute, R. (2018). IoT based greenhouse monitoring system. *International Journal for Research in Applied Science and Engineering Technology*, 6, 2084–2085. <https://doi.org/10.22214/ijraset.2018.4354>
11. Vimal, P. V., & Shivaprakasha, K. S. (2017). IoT based greenhouse environment monitoring and controlling system using Arduino platform. In *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)* (pp. 1514–1519). IEEE. <https://doi.org/10.1109/ICICICT1.2017.8342795>
12. Vimal, P. V., & Shivaprakasha, K. S. (2017). IoT based greenhouse environment monitoring and controlling system using Arduino platform. *IEEE International Conference on Intelligent Computing, Instrumentation and Control Technologies*, 1514–1519.
13. Qiang, G., & Ming, C. (2008). Research and design of web-based wireless sensor network management system for greenhouse. In *ICCEE*.
14. Rangan, K., & Vigneswaran, T. (2010). An embedded systems approach to monitor green house. In *Recent Advances in Space Technology Services and Climate Change (RSTSCC)* (pp. 61–65).
15. Attalla, D., & Tannfelt Wu, J. (2015). Automated greenhouse: Temperature and soil moisture control.
16. Gaikwad, A., Ghatge, A., Kumar, H., & Mudliar, K. (2016). Monitoring of smart greenhouse.
17. Hussain, R. H., Marhoon, A. F., & Rashid, M. T. (2013). Wireless monitor and control system for greenhouse. *International Journal of Computer Science and Mobile Computing*, 2, 69–87.
18. Jeyashree, K., & C. G. (2016). Monitor and control of environment for greenhouse using sensor networks. *International Journal of Advanced Research in Electronics and Communication Engineering*, 5.
19. Asolkar, P., & Bhadade, U. (2016). Analyzing and predicting the green house parameters of crops.

DESIGN AND IMPLEMENTATION OF A PRIVACY-PRESERVING IDENTITY MECHANISM FOR FILE STORAGE SERVERS

**Devayani Shailendra Singh Girase¹, Vaishnavi Sandip Jagtap¹,
Hiraman Jadhav^{*1}, Mayuri Rathi¹ and Ruchi Thakare¹**

¹Department of Information Technology,

²Department of Artificial Intelligence and Data Science,

³Department of Electronics and Telecommunication Engineering,

D. Y. Patil College of Engineering, Akurdi, Pune, Maharashtra, India

*Corresponding author E-mail: hiraman.jadhav123@gmail.com

Abstract

The concept 'server' is designed to serve information or services to multiple users over the network, which makes it, enable user profile management. Taken an example of storage server; system need to manage data or files from multiple users or belonging to multiple owners. Managing this ownership, systems normally separate the user files in different directory and records this directory information in index table. By the time system separated the files in different directory for its own management it unknowingly reveals the file owner information to system user who has direct access to the server directory structure. Main threat lies here only, that malicious user has access to every user's files. Malicious user attacks are not controllable or having no direct protection over it but it can be made difficult for the malicious user to get the ownership information about files by hiding the files original name and the ownership information by designing novel index table which has no record about files real identity and its location information. So, by the time of accessing the files its location and name information will be generated temperately using secure key.

Keywords: Server Security, File Ownership, Data Privacy, Secure Indexing, Access Control.

Introduction

Almost in every well know file storage server there is arrangement of separating the user files in user specific folder or directory which is normally named with user identification or user name. This make folders vulnerable and searchable if hacker is looking for specific user's files. Almost in every file storage server, it has a user management by allotting different folders to every user, but which creates a threat of user ownership identification, where by looking at server configuration one can identify the owner of the files. Proposed system is Invisible Ownership System. Where, no entity including file storage server having full information about files stored on the storage media. It simply means system will store all users' files on same drive or same folder. Now this will protect owner identity and file access threats. But for this user need to

replace complete file storage services on the entire network. For the proof of concept, we can rebuild file storage service and demonstrate it over intranet. Here, first client will pass the file-to-file storage service along with some security Key; file storage service will then rename the file with unique identification number or code which will be generated by user key. Once this new renamed file stored on the storage owner identity will be lost, hence system will be having no record or information about file owner detail along with file name. Every time while reading or accessing the file or file part system will calculate or generate the name of the file and access it.

Background

Looking toward the limitations, described in the literature review, it is necessary to have a novel universal procedure to overcome those limitations. As almost every big organization having web identity and runs on hosting or cloud environment, the organizational information becomes important aspect. Even after cloud provides industry a wide range of security and system free from malicious software virus attacks, still the data is not safe from malicious users having administrative rights. One having super user rights can access the data from cloud storage with wrong intentions. After going through different solutions to reduce malicious user attack we found keeping the data out of user reach will make data more secure than by any other way, but we can't forget the fact we have to keep the data somewhere. Hence the proposed system is designed by keeping these scenarios in consideration where we will distribute and secure the data at different location in order to hide original data directly from user. Here system will take smart decision on the basis of user request and split the data after successful encryption in to different blocks and store it on the multiple cloud storage, now malicious user can access the data but of no use. Another point of consideration is the safety of data in case of critical condition. The proposed system is less with capability to re calculate the data from rest of the storage if any one of the storage gets failure

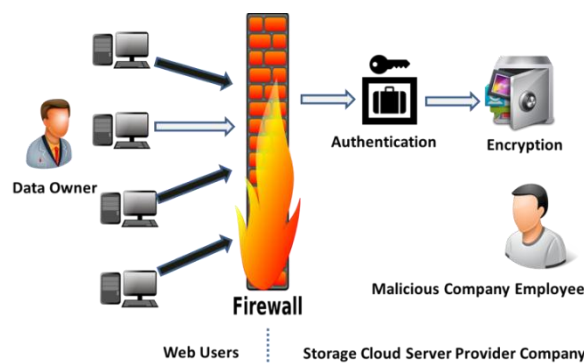


Figure 1: Security breaches

Different mechanism including authentication using credentials, defining file ownership in order to restrict the file access, firewalls, encryption mechanism are available to protect information. It is unfortunate that none of the stated method guaranteed the complete security of the data and even after providing high security the major threat to the data on public cloud is malicious user

figure 2.0 state the same issue where user can encrypt the information over cloud in secure environment but at the same time these storage locations are always accessible to the malicious user who is part of the public cloud service provider infrastructure. This thread is not easily traceable and no direct solution can be provided to this issue. D-HOM system mainly aims at hiding data and its ownership identity to make data inaccessible to malicious user which leads to securing the confidential data. Basic idea is to split file F in multiple parts $F1$, $F2$ and $F3$ then encrypting in to $F1e$, $F2e$ and $F3e$ respectively finally 'PUSH' it to public cloud storage.

Almost in every well know file storage server there is arrangement of separating the user files in user specific folder or directory which is normally named with user identification or user name. This makes folders vulnerable and searchable if hacker is looking for specific user's files. Take an example of file storage server, file storage service has a user management by allotting different folders to every user, but which creates a biggest threat of user ownership identification where by looking at service configuration one can identify the owner of the files. Part of the proposed system is Invisible Ownership System. Here no entity including file storage service has information about files stored on the storage. It simply means system will store all users' files on same drive or same folder.

Now this will protect owner identity and file access threats. But for this we need to replace complete file storage service access on the entire cloud service provider. For the proof of concept, we can rebuild file storage service and demonstrate it over intranet. Now how it will work, first client will pass the file-to-file storage service along with some Key, file storage service will then rename the file with unique identification number or key which is generated by user key and private key or something like that.

Once this new renamed file stored on the storage owner identity will be lost, no system will be there to identify, which file contains what and who is the owner of the file. See how useful it will be if we split then hide the identity of file. Every time while reading or accessing the file or file part system will calculate or generate the name of the file and access it.

Proposed System

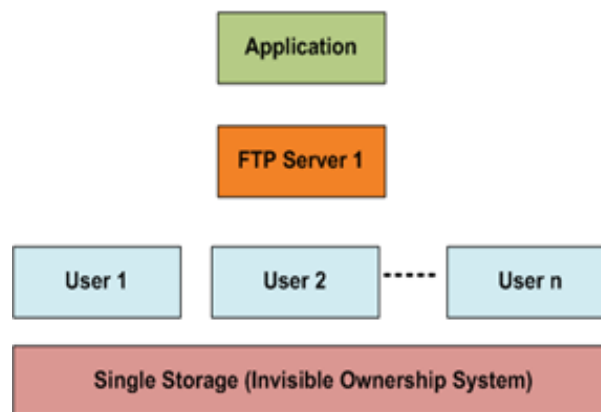


Figure 2: System protocol architecture

This present invention discloses system and method for secure file storage using hidden owner identity mechanism. Currently with the increase in the creation and use of digital data, enormous amount of data is created on file storage server which is maintained by third party vendor hence there is need to provide secure system which will store all user's data securely on system where only authorized user knows complete information.

Since file storage server stores all users data in various user directories, creates threat of user ownership recognition where system administrator attack is possible. We have developed secure file storage system where file owner information user is hidden hence nobody including user, administrator and file storage server has complete information about stored data. Our system reduces indexing overheads and cumbersome security key management.

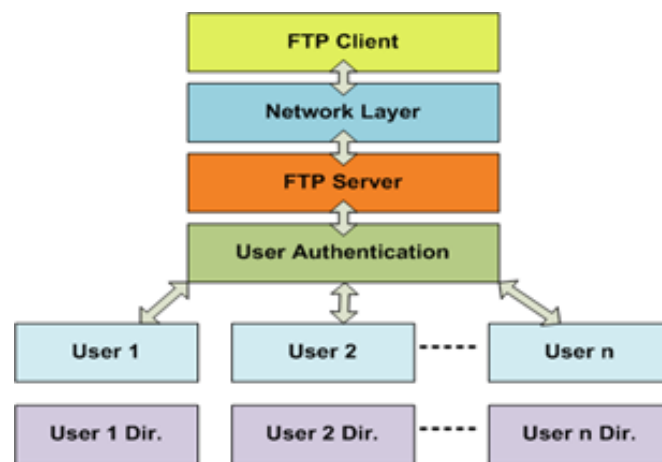


Figure 3: Existing FTP process

Invisible Ownership System

- Proposed New Method /Algorithm for improving FTP services
- Header Deletion function will remove ownership details of file chunks to provide OS level security
- Proposed system will dynamically create naming convention ambiguity by renaming each file
- New name is generated using Filename, Server name, user key and cloud private key

This present disclosure relates to the novel work carried for implementation of a Hidden Owner Identity mechanism for securely storing file on storage server. Moreover, it relates to unique, efficient and secure method for storing a file on any type of storage server either local or cloud-based server. Furthermore, present disclosure relates to computer program product consisting of computer -readable instructions stored on computer -readable storage media. These instructions being executed by computing system consisting of processing hardware to execute programs. File storage server is a server which stores various types of critical user data files and privacy -sensitive information using file systems therefore they are main targets for various types of security attacks.

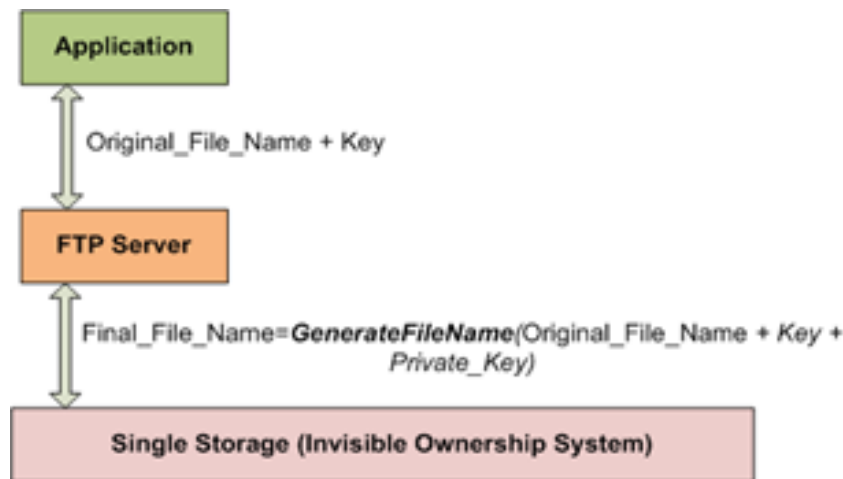


Figure 4: System process

Implementation

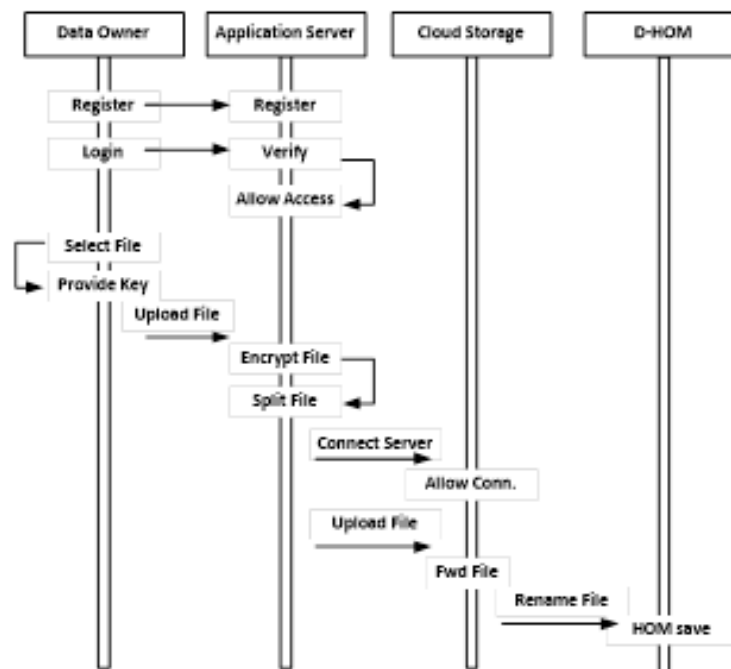


Figure 5: Process sequence diagram

The sequence diagram represent timeline over the file processing in our implementation. The first step is the registration of client over the application, generate the login credentials to make the request for the verification and authentication of client to start file processing over the developed environment. After getting all the credentials for file processing client is allow to start working over the environment. Client selects a file to be uploaded and generate a secure key to make a request to upload a file into the available cloud. The next step calls the Encryption and splitting API's to perform the secure distribution of file over the cloud. The timeline process calls also be called in reverse to get file receive request made satisfaction of the client.

Algorithm: Save file (HOM)

1. Get File $\rightarrow$ FL
 2. Get Key $\rightarrow$ K
 3. Generate GUID $\rightarrow$ G
 4. Rename FL=Enc(Gk)
 5. CreateFile(FP)
 6. Write FL $\rightarrow$ FP
 7. Save FP $\rightarrow$ CentralStorage
-

Conclusion

In presented proposed system we analysed different aspects of cloud computing and different file storage mechanism including local and distributed file system. Across the development we studied different available solutions and tried to find out what parameters can be considered and re-designed to enhance the performance or minimize the cost factor. We carried out the development of D-HOM (Distributed Hidden Ownership Mechanism) system which will enhance the file security by encrypting and distributing the file parts over multiple cloud and re-join when accessing. With simulation and implementation, we generated the performance result and resulting parameters are enhanced.

References

1. Bohli, J.-M., Gruschka, N., Jensen, M., Lo Iacono, L., & Marnau, N. (2013). Security and privacy enhancing multi-cloud architectures. *IEEE Transactions on Dependable and Secure Computing*.
2. Xiao, Z., & Xiao, Y. (2012). Security and privacy in cloud computing. *IEEE Communications Surveys & Tutorials*. Accepted for publication, March 2012.
3. Malik, A., & Nazir, M. M. (2012). Security framework for cloud computing environment. *Journal of Emerging Trends in Computing and Information Sciences*, 3(3).
4. Singhal, M., & Chandrasekhar, S. (2013). Collaboration in multicloud computing environments: Framework and security issues. IEEE Computer Society.
5. AlZain, M. A., Pardede, E., Soh, B., & Thom, J. A. (2012). Cloud computing security: From single to multi-clouds. In *Proceedings of the International Conference on System Sciences*.
6. Yang, K., Ren, X., Jia, X., Zhang, B., & Xie, R. (2013). DAC-MACS: Effective data access control for multi-authority cloud storage systems. IEEE.
7. Mell, P., & Grance, T. (2011). *The NIST definition of cloud computing*. National Institute of Standards and Technology, Technical Report.

8. Hwang, J.-J., & Chuang, H.-K. (2012). A business model for cloud computing based on a separate encryption and decryption service. National Science Council of Taiwan Government, IEEE.
9. Bohli, J.-M., Jensen, M., Gruschka, N., Schwenk, J., & Lo Iacono, L. (2011). Security prospects through cloud computing by adopting multiple clouds. In *Proceedings of the IEEE Fourth International Conference on Cloud Computing (CLOUD)*.
10. Jensen, M., Schwenk, J., Gruschka, N., & Lo Iacono, L. (2009). On technical security issues in cloud computing. In *Proceedings of the IEEE International Conference on Cloud Computing (CLOUD-II)*.
11. Yang, K., & Jia, X. (2012). Attribute-based access control for multi-authority systems in cloud storage. In *Proceedings of the 2012 32nd IEEE International Conference on Distributed Computing Systems*. IEEE.
12. AlZain, M. A., Soh, B., & Pardede, E. (2011). MCDB: Using multi-clouds to ensure security in cloud computing. In *Proceedings of the 2011 Ninth IEEE International Conference on Dependable, Autonomic and Secure Computing*. IEEE.
13. Jeeva Rathanam, S. G., & Sumalatha, M. R. (2012). PDDS—Improving cloud data storage security using data partitioning technique. IEEE.
14. Mandal, A. K., & Tiwari, A. (2012). Performance evaluation of cryptographic algorithms: DES and AES. In *Proceedings of the 2012 IEEE Students' Conference on Electrical, Electronics and Computer Science*. IEEE.
15. J. D., Assistant Professor, Ramkumar, P., Systems Engineer, & Kadhivelu, D. (2010). Preserving privacy through data control in a cloud computing architecture using discretion algorithm. In *Proceedings of the Third International Conference on Emerging Trends in Engineering and Technology*. IEEE.
16. Kumar, P., & Kumar, L. (2013). Security threats to cloud computing. *International Journal of IT, Engineering and Applied Sciences Research (IJIEASR)*, 2(1).

Artificial Intelligence and Beyond: Data, Computing and Smart Systems (ISBN: 978-81-69967-09-9)

About Editors



Dr. Med Ram Verma is working as Principal Scientist and Head at ICAR-Indian Agricultural Statistics Research Institute, Pusa, New Delhi. He completed his M.Sc. in Agricultural Statistics in 1998 from Dr. Y.S. Parmar University of Horticulture and Forestry, Nauni, Solan, and Ph.D. in Statistics in 2005 from Dr. B.R. Ambedkar University, Agra. He joined the Research Service in 2003 and served as Senior Scientist at ICAR-Indian Veterinary Research Institute, Izatnagar, Bareilly, from December 2009 to July 2023. He has guided four Ph.D. and 10 M.V.Sc. students in Biostatistics and published 260 research papers. He has authored/edited 11 books. He received the Best Teacher Award for 2014–15 and the Bharat Ratna Dr. C. Subramaniam Award for Outstanding Teachers 2018. He is a Fellow of the Indian Society of Agricultural Statistics, elected member of the International Statistical Institute, Netherlands, and serves as an editorial board member.



Dr. K Pushpanjali Patra (M.Sc. Physics, CSIR-NET (JRF), GATE, Ph.D., IIT Guwahati) is an academican and researcher in Physics, with teaching, research, and publication experience. She serves as an Assistant Professor, researching nanomaterials and magnetic systems. She has published more than 19 SCIE-indexed research papers with citation impact. She has served as a reviewer for international journals and editor of books. Her research expertise includes low-temperature magnetic characterization, spin glass systems, magnetocaloric effect, tuneable exchange bias, negative magnetization, and advanced material analysis. She has participated in conferences on condensed matter physics and materials science and taught at UG/PG levels. Her academic expertise covers Quantum Computation, Semiconductor Physics, Electromagnetic Theory, Quantum Mechanics, Classical Mechanics, Thermodynamics, Mathematical Physics, and programming languages.



Dr. Renuka P. Chopde is an accomplished academican and emerging leader in Chemistry, serving as Assistant Professor in the Department of Chemistry at S. Chandra Mahila Mahavidyalaya, Sakoli, Bhandara. With over 14 years of research and 11 years of teaching experience, she specializes in Organic Chemistry, Polymer Chemistry, and advanced materials science. Her doctoral research focused on the synthesis and characterization of novel polymeric resins and their applications. She has published numerous research papers on ion-exchange and nanostructured materials and authored chemistry textbooks. Her contributions extend to university-level academic responsibilities, academic administration, and institutional development. She is committed to quality education, scientific inquiry, research innovation, women's empowerment in science, and holistic student development through teaching, research, and scholarly activities.



Er. Harshit Gupta is an academican and researcher with over 12 years of experience in education, banking, and corporate sectors. His academic expertise spans Computer Science and Engineering, Information Technology, Education, and Management. He serves as Head, Department of Computer Application and Programme Coordinator of M.Tech (CSE, AI-ML & DS) at Rajshree Institute of Management & Technology, Bareilly, affiliated with AKTU, Lucknow. He holds engineering degrees with Honors from GBTU/AKTU and degrees with Gold Medals from MJPRU, Bareilly. He has qualified UGC-NET and holds MCP and CCNA certifications. His research interests include Artificial Intelligence, Machine Learning, IoT, Prompt Engineering, Blockchain, Big Data, Data Analytics, Pattern Recognition, and Python. He has published over 200 reputed research papers, books, and book chapters.

