Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X Volume 2, Issue 5, 2024

FEDERATED LEARNING FOR THREAT INTELLIGENCE SHARING: A PRIVACY-PRESERVING COLLABORATIVE DEFENSE MODEL

Ahmad Bacha[*1], Hijab Sehar[2], Suhaib Naseem[3], Muhammad Ismaeel Khan[4]

[*1]*Washington University of Science and Technology, Alexandria, Virginia*
[2]*RiphahSchool of Computing and Innovation, Lahore*
[3]*FiveRivers Technologies, 26-F, Gulberg 2, Main Market, Lahore, Pakistan.*
[4]*School of Information Technology, Washington University of Science and Technology, USA*

DOI:https://doi.org/10.5281/zenodo.17657093

Keywords
Federated Learning, Threat Intelligence, Privacy-Preserving AI, Cross-Border Cybersecurity, Homomorphic Encryption, Data Sovereignty

Article History
Received: 05 December 2024
Accepted : 15 December2024
Published : 31 December2024

Corresponding Author: *
Ahmad Bacha

Abstract
The emergence of advanced, cross-national cyber threats is a reason that promotes the sharing of threat intelligence, which is often hampered due to the law of privacy, security interests, and political conflicts. Older centralized systems in which sensitive information is aggregated in one repository are becoming unsustainable. Federated Learning (FL) provides a paradigm shift and allows jointly training machine learning models by providing no raw data to any of the entities. This decentralized model lets the cross-border organizations enjoy a collective defense model without compromising their confidential information rights. In this paper, the design, implementation, and challenges of FL systems in achieving cybersecurity are discussed. It outlines architectural designs, privacy protection mechanisms, such as differential privacy and homomorphic encryption, and gives financial, critical infrastructure, and national security applications. The outcomes of the simulations reveal a reasonable privacy-utility trade-off and the paper ends with a roadmap to future research and real-life implementation, in which FL can serve as the foundation of contemporary, privacy-conscious cyber defense.

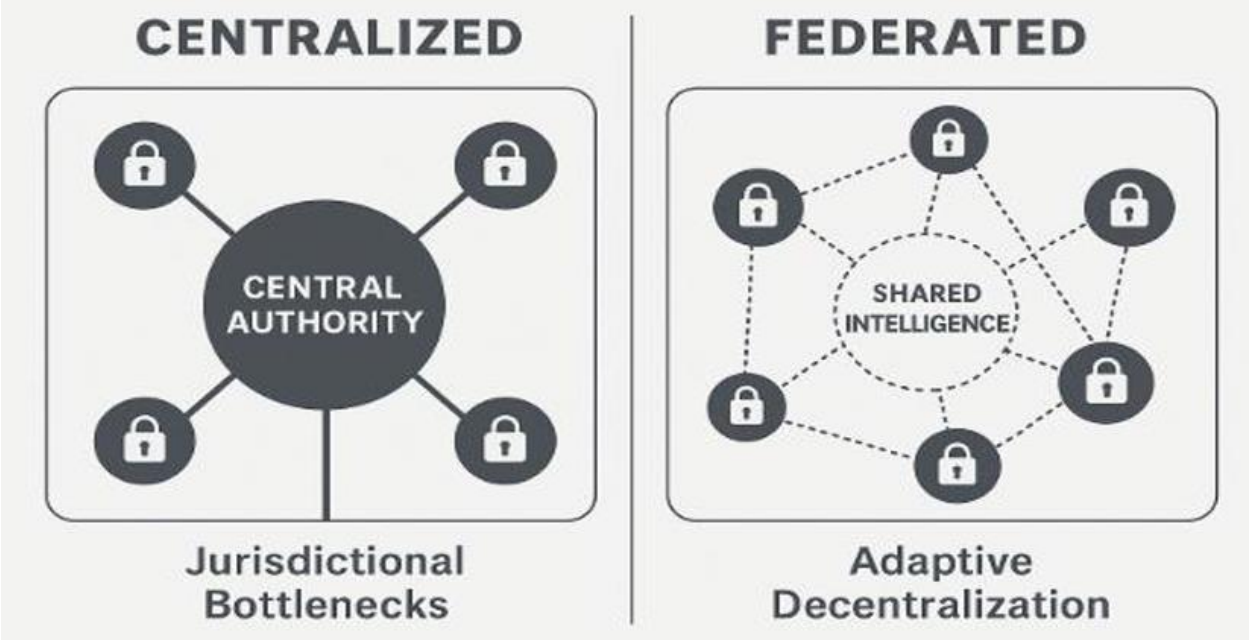
INTRODUCTION

The cybersecurity environment at the global level is characterized by more complex and multi-national attacks on critical infrastructure[1]. At the same time, data sovereignty laws and geopolitical distrust have resulted in a "cyber-fragmentation," that hinders the exchange of crucial parts of threat intelligence. The centralised models that consolidate data are politically and legally problematic, which prompts a very pressing necessity of the decentralised model [2].

Federated Learning (FL) is a way out of this impasse. It is a machine learning model in which a global model is trained on a number of decentralized clients with local samples of data. Rather than communicating raw data, the participants learn models in advance and only distribute model updates, which are used to refine the global model. This means that raw data does not move outside of the system, therefore maintaining privacy and yet adhering to domestic laws [3].

This paper examines the application of FL to cybersecurity threat intelligence[4]. We explore its architecture, the privacy-preserving mechanisms that secure it, and its performance in simulated and real-world scenarios[5]. The objective is to provide a comprehensive overview of FL as a

viable framework for building a collaborative, yet privacy-conscious[6], global cyber defense network[7],[8].



Centralized vs. Federated collaboration models, highlighting jurisdictional bottlenecks versus adaptive decentralization.

Figure 1: Centralized vs. Federated collaboration models, highlighting jurisdictional bottlenecks versus adaptive decentralization.

2. FEDERATED LEARNING FUNDAMENTALS & CHALLENGES IN CYBERSECURITY

2.1 FL Concepts and Taxonomy

FL operates on the principle of decentralized data and centralized model aggregation. Key variants include:

- **Horizontal FL:** Used when participants share the same feature space[9] but have different data subjects (e.g., multiple hospitals detecting the same malware) .
- **Vertical FL:** Applicable when different entities hold different features for the same data subjects (e.g., a bank and a telecom company collaborating on fraud detection) .
- **Cross-Silo FL:** Involves a small number of reliable organizations (e.g., national CSIRTs, large corporations) with substantial computational resources .
- **Cross-Device FL:** Involves a large number of resource-constrained devices (e.g., IoT sensors, mobile phones) .

For cross-border threat intelligence, Cross-Silo FL is the most relevant and stable architecture.

2.2 Challenges of Traditional Threat Intelligence Sharing

Traditional centralized AI models for cybersecurity[10],[11], embedded in SIEMs and IDS/IPS, face critical limitations in a global context:

- **Data Sovereignty:** Laws in Russia, India, and the EU prohibit certain data from leaving national borders .
- **Privacy Regulations:** GDPR, CCPA, and others make sharing even anonymized data a compliance risk .
- **High-Value Targets:** Centralized data repositories are attractive targets for attackers.
- **Operational Limitations:** They perform poorly in air-gapped or low-connectivity environments and lack real-time adaptability .

Table 1: Conventional vs. Federated AI Cybersecurity Frameworks

Criterion	Conventional Cybersecurity	Federated AI Frameworks
Data Privacy Compliance	Low	High
Operational Resilience	Low (single point of failure)	High (distributed)
Performance in Isolated Zones	Poor	Good
Legal Interoperability	Limited	Strong
Scalability	Constrained	Elastic
Real-Time Adaptability	Limited	High

3. SYSTEM ARCHITECTURE AND DESIGN

3.1 Federated Architecture for Threat Detection

A robust FL architecture for cybersecurity[12] involves multiple independent clients (e.g., CSIRTs of different nations) and a central aggregator .

1. **Local Training:** Each client trains a model on its local threat data (network flows, endpoint logs).
2. **Secure Transmission:** Encrypted model updates are sent to a federated server.
3. **Secure Aggregation:** The server aggregates these updates (e.g., using FedAvg) to create an improved global model.
4. **Global Model Distribution:** The updated model is sent back to all clients.

A secure relay layer using Homomorphic Encryption (HE) or Secure Multi-Party Computation (SMPC) can be added to protect updates in transit. A policy control plane governs trust levels and privacy thresholds. This flow is depicted in **Figure 2**.

3.2 Threat Signals and Model Types

FL models in cybersecurity are trained on diverse, privacy-conscious signals:

- **Network Flow Telemetry:** IP headers, port numbers, session data.
- **Endpoint Security Logs:** File hashes, process trees, behavioral scores.
- **DNS Telemetry:** Lookups to malicious domains.
- **Anomaly Scores:** Outputs from local unsupervised models.

Common model types and their objectives include:

- **Intrusion Detection:** RNNs/LSTMs for temporal sequence analysis.
- **Malware Classification:** CNNs (for image-based binary analysis) or tree-based ensembles (for tabular log data).
- **Zero-Day Detection:** Unsupervised models like autoencoders for anomaly detection.

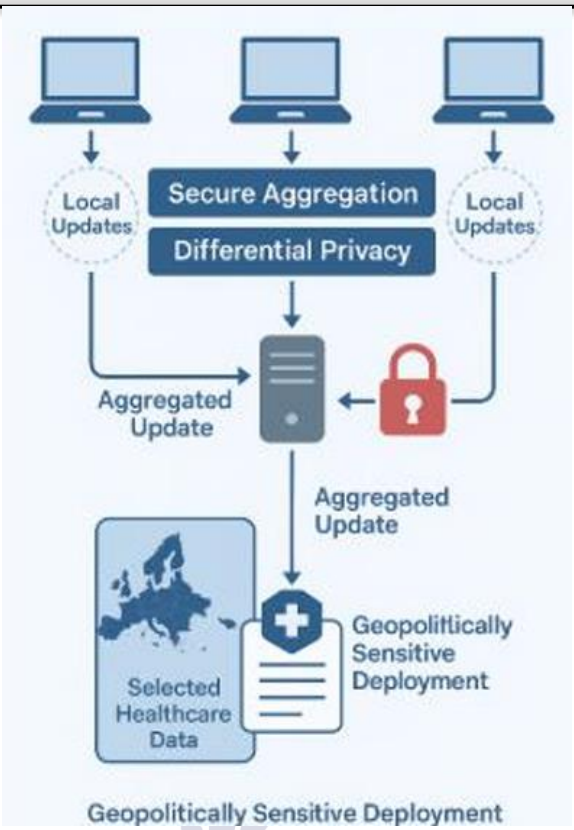


Figure 2: Secure FL architecture with cryptographic protections for model updates

4. PRIVACY-PRESERVING MECHANISMS

To mitigate risks like model inversion and membership inference attacks, FL systems employ several key techniques, compared in **Table 2**.

4.1 Differential Privacy (DP)

DP adds calibrated mathematical noise to model updates before they are shared. This ensures that the contribution of any single data point is obscured, providing formal, auditable privacy guarantees bounded by a parameter (ε). The key challenge is balancing the noise (privacy) with model accuracy (utility) [15] [16].

4.2 Homomorphic Encryption (HE) & Secure Multi-Party Computation (SMPC)

- **HE** allows computations to be performed directly on encrypted data. Clients send encrypted updates, and the server aggregates them without decryption, revealing only the final combined result [17] [18].
- **SMPC** distributes trust by splitting a client's update into secret shares sent to multiple aggregators. The global model is reconstructed without any single party seeing a complete update [19] [20].

While both provide strong security, they introduce significant computational and communication overhead.

4.3 Secure Aggregation

This is a foundational protocol that ensures the aggregation server only ever sees the sum of client updates, not any individual contribution. It often uses cryptographic masking techniques and is resilient to client dropouts, making it essential for cross-jurisdictional collaboration [21] [22].

Table 2: Comparison of Privacy-Preserving Techniques

Technique	Privacy Guarantee	Compute Overhead	Impact on Accuracy	Best Use Case

Differential Privacy	Moderate	Low-Medium	Medium	Large-scale detection with legal audit
Homomorphic Encryption	Strong	High	Low	Defense, healthcare
SMPC	Strong	Medium-High	Low-Medium	Multi-agency coalitions
Secure Aggregation	Strong	Medium	Negligible	Cross-border collaboration

5. USE CASES AND SIMULATION RESULTS

5.1 Real-World Use Cases

1. **Cross-National Cyber Defense:** NATO-aligned CSIRTs can use FL to detect Advanced Persistent Threats (APTs) that cross borders. Each national node trains on local forensic data, and encrypted updates create a global model that identifies campaign patterns without sharing raw, classified logs.
2. **Financial Sector Collaboration:** Banks across different jurisdictions (e.g., London, New York, Singapore) can collaboratively train fraud and malware detection models. This helps identify coordinated attacks targeting the global financial system while adhering to strict client confidentiality laws like GLBA.
3. **Critical Infrastructure Protection:** Energy grid operators in different countries can use FL to detect anomalies in Industrial Control Systems (ICS) and SCADA data. This enables a coordinated defense against cyber-physical attacks without exposing sensitive operational blueprints [23].

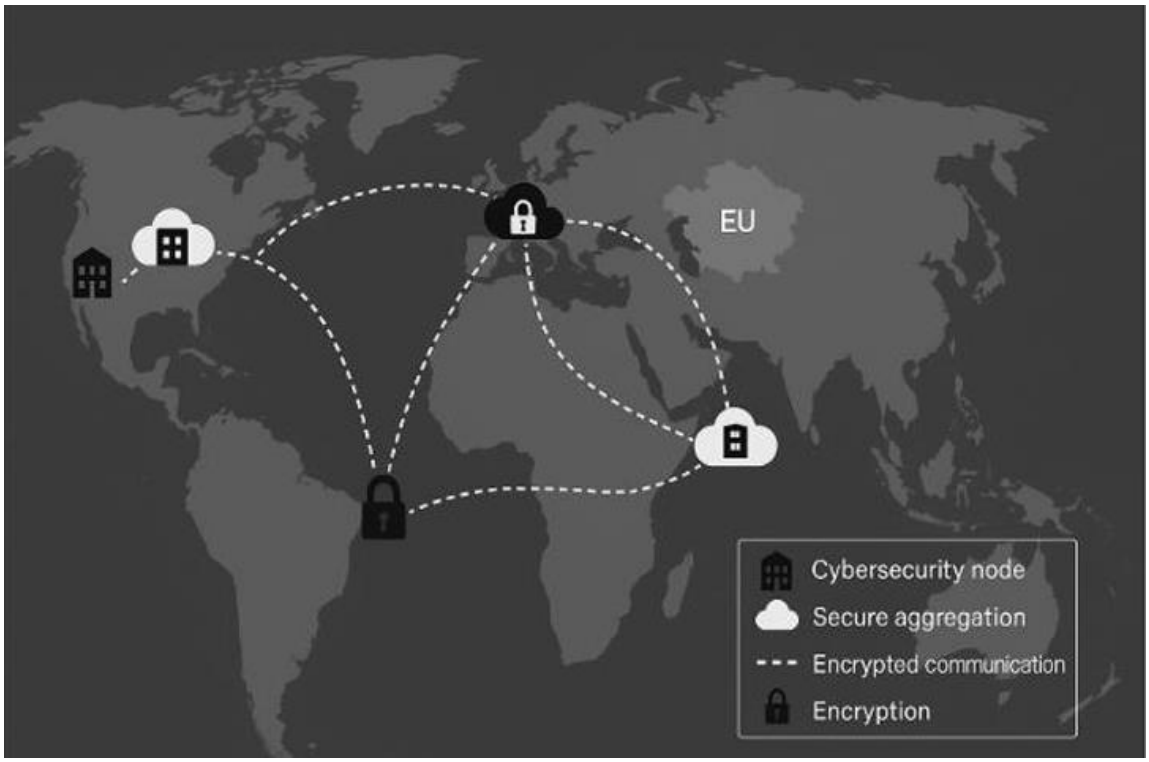


Figure 3: Global map showing federated nodes for CSIRTs, financial institutions, and infrastructure agencies.

5.2 Simulation Framework and Results

A simulation was conducted with 12 synthetic organizational nodes across different regions, generating over 600,000 log entries with injected attacks. The evaluation compared several configurations:

- Centralized Learning
- Basic FL (no extra security)
- FL with Differential Privacy (FL+DP)
- FL with SMPC (FL+SMPC)
- FL with both DP and SMPC (FL+DP+SMPC)

Key Findings (Summarized in Table 3):

- **Centralized learning** had the highest accuracy (F1: 92.1%) but the worst privacy (Leakage: 0.25).
- **Basic FL** maintained high accuracy (F1: 90.5%) with improved privacy (Leakage: 0.11).
- **FL+DP** reduced leakage to 0.07 but incurred a performance cost (F1: 84.2%).
- **FL+SMPC** offered a good balance, with strong privacy (Leakage: 0.03) and good accuracy (F1: 88.9%).
- **FL+DP+SMPC** provided the highest security (Leakage: <0.01) with acceptable accuracy (F1: 83.1%) for high-risk sectors.

These results confirm that FL with appropriate safeguards can provide a superior privacy-utility balance for global threat intelligence.

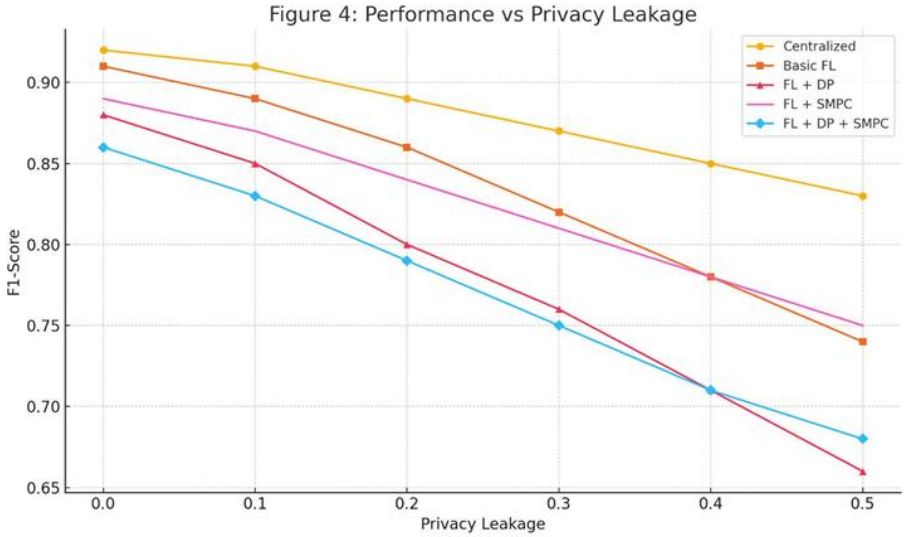


Figure 4: Performance curves comparing F1-score and privacy leakage across different FL configurations.

Table 3: Quantitative Performance and Privacy Summary

Configuration	F1-Score	ROC-AUC	Privacy Leakage	Latency (ms/round)
Centralized	92.1%	0.964	0.25	105
Basic FL	90.5%	0.956	0.11	126
FL + DP	84.2%	0.902	0.07	130
FL + SMPC	88.9%	0.937	0.03	157
FL + DP + SMPC	83.1%	0.895	0.01	172

6. THREATS, DEFENSES, AND IMPLEMENTATION CHALLENGES

6.1 Threat Model

FL introduces unique attack vectors:

- **Model Poisoning:** Malicious clients send manipulated updates to corrupt the global model.
- **Privacy Inference:** Adversaries reverse-engineer training data from shared gradients (model inversion) or deduce data membership.
- **Collusion:** Multiple malicious nodes collaborate to attack the system.

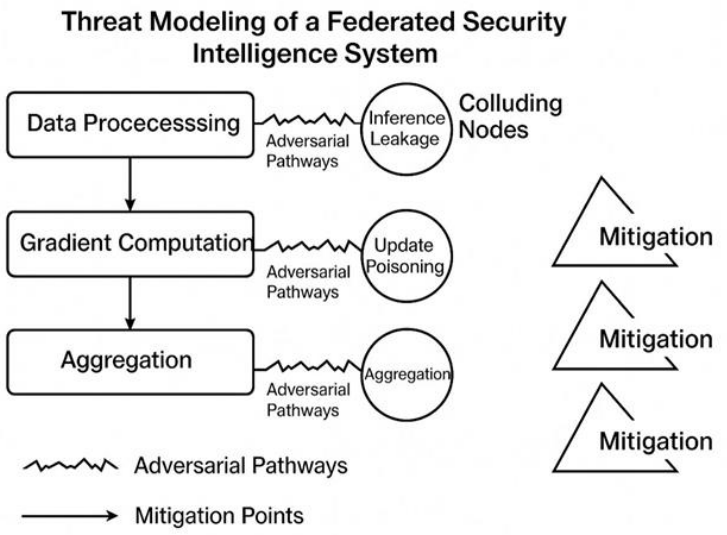


Figure 5: Threat model diagram showing adversarial pathways and defense points in an FL system.

6.2 Defense Strategies

A multi-layered defense is required:

1. **Trust Calibration:** Dynamically weight client contributions based on historical reliability.
2. **Anomaly Detection at Aggregator:** Monitor updates for statistical outliers to detect poisoning.
3. **Robust Aggregation Algorithms:** Use algorithms like FedProx that are resilient to non-IID data and malicious updates.
4. **Scheduled Retraining & Rollback:** Maintain model versions to revert to a safe state if corruption is detected.

6.3 Implementation Challenges

- **Performance Overhead:** Encryption (HE, SMPC) and secure aggregation increase latency and computational cost.
- **Synchronization:** Coordinating updates across global time zones and networks can lead to model drift.
- **Organizational Trust:** Achieving buy-in requires transparent governance, legal agreements (MOUs), and auditability.
- **Compliance and Interoperability:** FL systems must integrate with existing SOC tools (SIEM, SOAR) and comply with standards like NIST and ISO 27001.

7. FUTURE DIRECTIONS & CONCLUSION

7.1 Roadmap

The future of FL in cybersecurity lies in:

1. **Edge Deployment:** Embedding FL clients directly into firewalls and IDS for real-time, low-latency threat detection.
2. **Zero-Day and APT Detection:** Using meta-learning techniques in FL to generalize from local anomalies to identify novel, global threats.
3. **Decentralized Trust Models:** Leveraging blockchain and smart contracts for automated, transparent governance and node compliance.
4. **Federated AI Alliances:** Establishing neutral, international FL consortia under bodies like the UN to foster cyber peacebuilding among adversarial states.

7.2 Conclusion

Federated Learning is a game changer in terms of collaboration in cybersecurity in the disjointed world. It facilitates the exchange of knowledge without data transfer, overcoming the challenges of data sovereignty and privacy issues and the major concern of geopolitical mistrust directly. Although technical challenges in the areas of performance and security still exist, the concept of integrating both: differential privacy and homomorphic encryption, along with strong

aggregation protocols gives a clear way ahead. To achieve all the objectives of FL, researchers, industry leaders, and policymakers must work together to implement uniform treatment protocols, build trust systems, and encourage implementation. Federated Learning is not only a technical advancement, but a needed move towards a more collaborative, resilient, and privacy-sensitive future of the world in cyber defense.

REFERENCES

[1] Khan, Muhammad Ismaeel, Hassan Tahir, Aftab Arif, Md Ismail Jobiullah, Ali Raza A. Khan, Sakera Begum, and Ihtasham Hafeez. "Enhancing IoT Security: A Lightweight Cloning Approach for RFID/NFC Access Control Systems." Cuestiones de Fisioterapia 52, no. 2 (2023): 231-248.

[2] D. Usynin et al., "Adversarial interference and its mitigations in privacy-preserving collaborative machine learning," Nat. Mach. Intell., vol. 3, no. 9, pp. 749–758, 2021.

[3] X. Yin, Y. Zhu, and J. Hu, "A Comprehensive Survey of Privacy-preserving Federated Learning: A Taxonomy, Review, and Future Directions," ACM Comput. Surv., vol. 54, no. 6, pp. 1–36, July 2022, doi: 10.1145/3460427.

[4] Khan, Muhammad Ismaeel, Aftab Arif, and Ali Raza A. Khan. "The most recent advances and uses of AI in cybersecurity." BULLET: Jurnal Multidisiplin Ilmu 3, no. 4 (2024): 566-578.

[5] Arif, Aftab, Fadia Shah, Muhammad Ismaeel Khan, Ali Raza A. Khan, Aftab Hussain Tabasam, and Abdul Latif. "Anomaly detection in IoHT using deep learning: Enhancing wearable medical device security." Migration Letters 20, no. S12 (2023): 1992-2006.

[6] Sikander Niaz et al., "AI for Inclusive Educational Governance and Digital Equity Examining the Impact of AI Adoption and Open Data on Community Trust and Policy Effectiveness," Contemporary Journal of Social Science Review 2, no. 4 (2024): 2560, https://doi.org/10.63878/cjssr.v2i04.1502.

[7] Akbar, Z., Hassaan, A., Jamshaid, M. M., Siddique, M. N., & Niaz, S. (2023). Leveraging Data and Artificial Intelligence for Sustained Competitive Advantage in Firms and Organizations. Journal of Innovative Computing and Emerging Technologies, 3(1).

[8] Muhammad Mudaber Jamshaid, Ahmed Hassaan, Zeeshan Akbar, Muhammad Nouman Siddique, & Sikander Niaz. (2024). IMPACT OF ARTIFICIAL INTELLIGENCE ON WORKFORCE DEVELOPMENT: ADAPTING SKILLS, TRAINING MODELS, AND EMPLOYEE WELL-BEING FOR THE FUTURE OF WORK. Spectrum of Engineering Sciences, 2(1). Retrieved from https://thesesjournal.com/index.php/1/article/view/1417

[9] Hassaan, A., Jamshaid, M. M., Siddique, M. N., Akbar, Z., & Niaz, S. (2023). ETHICAL ANALYTICS & DIGITAL TRANSFORMATION IN THE AGE OF AI: EMBEDDING PRIVACY, FAIRNESS, AND TRANSPARENCY TO DRIVE INNOVATION AND STAKEHOLDER TRUST. Contemporary Journal of Social Science Review, 1(04), 1-18.

[10] Shah, Heta Hemang, and Ahmad Bacha. "Leveraging AI and Machine Learning to Predict and Prevent Sudden Cardiac Arrest in High-Risk Populations." Global Journal of Universal Studies 1, no. 2 (2024): 87-107.

[11] Bacha, Ahmad, and Heta Hemang Shah. "AI-Enhanced Liquid Biopsy: Advancements in Early Detection and Monitoring of Cancer through Blood-based Markers." Global Journal of Universal Studies 1, no. 2 (2024): 68-86.

[12] Aziz, Rimsha, Aneela Mehmood, Asma Tariq, Fawad Nasim, Umar Farooq, Syed Asad Ali Naqvi, and Hamayun Khan. "Critical Evaluation of Data Privacy and Security Threats: An Intelligent Federated Learning-based Intrusion Detection System Poisoning Attack and Defense for Cyber-Physical Systems its Issues and Challenges Related to Privacy and Security in IoT." The Asian Bulletin of Big Data Management 5, no. 1 (2025): 73-84.

[13] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Federated learning for data privacy preservation in vehicular cyber-physical systems," IEEE Netw., vol. 34, no. 3, pp. 50–56, 2020.

[14] G. Long, T. Shen, Y. Tan, L. Gerrard, A. Clarke, and J. Jiang, "Federated Learning for Privacy-Preserving Open Innovation Future on Digital Health," in Humanity Driven AI, F. Chen and J. Zhou, Eds., Cham: Springer International Publishing, 2022, pp. 113–133. doi: 10.1007/978-3-030-72188-6_6.

[15] G. Abad, S. Picek, V. J. Ramírez-Durán, and A. Urbieta, "On the Security & Privacy in Federated Learning," Mar. 16, 2022, arXiv: arXiv:2112.05423. doi: 10.48550/arXiv.2112.05423.

[16] M. Akter, N. Moustafa, T. Lynar, and I. Razzak, "Edge intelligence: Federated learning-based privacy protection framework for smart healthcare systems," IEEE J. Biomed. Health Inform., vol. 26, no. 12, pp. 5805–5816, 2022.

[17] J. Domingo-Ferrer, A. Blanco-Justicia, J. Manjón, and D. Sánchez, "Secure and privacy-preserving federated learning via co-utility," IEEE Internet Things J., vol. 9, no. 5, pp. 3988–4000, 2021.

[18] M. Hao, H. Li, X. Luo, G. Xu, H. Yang, and S. Liu, "Efficient and privacy-enhanced federated learning for industrial artificial intelligence," IEEE Trans. Ind. Inform., vol. 16, no. 10, pp. 6532–6542, 2019.

[19] G. P. Rusum and K. K. Pappula, "Federated Learning in Practice: Building Collaborative Models While Preserving Privacy," Int. J. Emerg. Res. Eng. Technol., vol. 3, no. 2, pp. 79–88, 2022.

[20] P. F. Saura, J. F. M. Gil, J. B. Bernabé, and A. Skarmeta, "Privacy-Preserving Cyber Threat Information Sharing Leveraging FL-Based Intrusion Detection in the Financial Sector," in Digital Sovereignty in Cyber Security: New Challenges in Future Vision, vol. 1807, A. Skarmeta, D. Canavese, A. Lioy, and S. Matheu, Eds., in Communications in Computer and Information Science, vol. 1807. , Cham: Springer Nature Switzerland, 2023, pp. 50–64. doi: 10.1007/978-3-031-36096-1_4.

[21] Y. Chen et al., "DS2PM: A data-sharing privacy protection model based on blockchain and federated learning," IEEE Internet Things J., vol. 10, no. 14, pp. 12112–12125, 2021.

[22] Y. Li, Y. Zhou, A. Jolfaei, D. Yu, G. Xu, and X. Zheng, "Privacy-preserving federated learning framework based on chained secure multiparty computing," IEEE Internet Things J., vol. 8, no. 8, pp. 6178–6186, 2020.

[23] M. Ali, F. Naeem, M. Tariq, and G. Kaddoum, "Federated learning for privacy preservation in smart healthcare systems: A comprehensive survey," IEEE J. Biomed. Health Inform., vol. 27, no. 2, pp. 778–789, 2022.