World Journal of Advanced Engineering Technology and Sciences

eISSN: 2582-8266

Cross Ref DOI: 10.30574/wjaets

Journal homepage: https://wjaets.com/

(Research Article)

Privacy and data protection challenges in industry 4.0: An AI-driven perspective

Jauhari Tanisha [2, 4, *], Pillai Adithya Rajesh [2], Roy Gaurpriya Singh [3], Koshy Adhip [2], Kothari Stuti [3] and D. Ajitha [1]

[1] *Department of Software Systems, School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology Vellore, Tamil Nadu, India.*
[2] *School of Computer Science and Engineering (SCOPE), Vellore Institute of Technology Vellore Tamil Nadu, India.*
[3] *School of Information Technology (SCORE), Vellore Institute of Technology Vellore Tamil Nadu, India.*
[4] *School of Computer Science and Engineering, SCOPE, Tamil Nadu, India.*

World Journal of Advanced Engineering Technology and Sciences, 2024, 12(02), 064–089

Publication history: Received on 31 May 2024; revised on 07 July 2024; accepted on 09 July 2024

Article DOI: https://doi.org/10.30574/wjaets.2024.12.2.0287

Abstract

In Industry 4.0, data is enormously exchanged through wireless devices. Ensuring data privacy and protection is vital. The proposed review paper explores how AI techniques can safeguard sensitive information and ensure compliance. It explores fundamental technologies such as cyber-physical systems and the complex application of AI in analytics and predictive maintenance. The issues with data security are then emphasized and privacy concerns resulting from human-machine interaction are shown. Regulatory frameworks that direct enterprises are touted as essential defenses, coupled with AI-powered solutions and privacy-preserving tactics. Examples from everyday life highlight the constant battle for equilibrium. The review continues with a look ahead to future developments in interdisciplinary research and ethical issues that will influence Industry 4.0's responsible growth. In essence, this paper synthesizes a nuanced understanding of the sophisticated challenges surrounding privacy and data protection within Industry 4.0, underscoring the pivotal role of AI as a custodian of sensitive information and offering an indispensable resource for professionals, policymakers, and researchers navigating the intricate and evolving terrain of Industry 4.0 with technical precision and ethical responsibility.

Keywords: Privacy Protection; Data Protection; Industry 4.0; AI.

1. Introduction

The dawn of Industry 4.0 represents a pivotal moment in the evolution of manufacturing as it ushers in the fourth industrial revolution. Industry 4.0 follows the footsteps of the first industrial revolution, marked by water and steam power, and the subsequent industrial revolutions, characterized by mass production assembly lines, and the digital revolution fueled by computers and information technology. What sets Industry 4.0 apart is its emphasis on automation, smart machines, and data-centric operations.

At the heart of Industry 4.0 lie several key technologies: Cloud computing, regarded as the backbone of Industry 4.0, provides the infrastructure for vast data storage, analysis, and supply chain integration. Artificial intelligence (AI) and machine learning step in with their powerful algorithms, driving automation, predictability, and optimization across manufacturing processes. Edge computing, minimized latency, enabling real-time data analysis in Industry 4.0 where cybersecurity is crucial for safeguarding interconnected equipment. Lastly, Digital Twins, virtual replicas of processes and supply chain, are pivotal for troubleshooting and data driven decision making. Integrating Information Technology (IT) & Operational Technology (OT) ensures seamless communication across factory assets and enterprise software [1].

[*] Corresponding author: Jauhari Tanisha.

Industry 4.0 delivers efficiency, predictive maintenance and transparent supply chains, aiming for lights out manufacturing. AI is pivotal, enhancing this transformation across sectors. Exploring AI's integration, applications, challenges, and its role in smart factories and decision-making. Five design principles drive this holistic integration for agile, data-driven ecosystems in diverse industries like healthcare, logistics and agriculture [2].

- Interoperability: Enables swift supply chain decision via seamless device data exchange.
- Real-time Data: Involves data analysis for process optimization and better product quality.
- Virtualization: Creates virtual models to simulate, maintain & prototype physical systems.
- Decentralization: Distributes decision-making in manufacturing, enhancing adaptability.
- Modularity: Enables swift system reconfiguration, reducing downtime for rapid adaptations.

Industry 4.0 aims to enhance manufacturing competitiveness but reveals challenges in managing data demands. It responds by leveraging extensive machine-to-machine communication and IoT, reshaping traditional practices into modern, intelligent systems. The integration of Industry 4.0 with AI drives unprecedented innovation but this rapid device proliferation also raises complex privacy and data protection issues to ensure authenticity and security of industrial data.

Organization of this paper: The paper's structure encompasses 13 sections focusing on various aspects of Industry 4.0 and AI within the context of privacy and data. Section 2 provides an overview of privacy, data, and artificial intelligence in Industry 4.0, while Section 3 outlines the research methodology for accuracy. Section 4 analyzes key challenges in data security and privacy, and Section 5 delves into privacy-preserving techniques from an AI perspective. Additionally, Section 6 explores regulatory frameworks and compliance as defense measures, followed by Section 7 addressing the limitations of these regulations. Real-world scenarios from field research are discussed in Section 8, while Section 9 explores cross-disciplinary approaches. Section 10 extrapolates future trends and scopes based on previous solutions, followed by an examination of ethical considerations in Section 11. Section 12 focuses on the limitations of the paper, and finally, Section 13 concludes the findings and discussions.

1.1. Privacy concerns in Industry 4.0

Smart factories have ushered in a new wave of interconnected devices that facilitate real-time data sharing over wireless connections, which raise critical data privacy issues. It is to be ensured that data sharing transpires only among trusted entities while upholding the pillars of privacy and security. This is achieved by implementing robust access control and user management mechanisms. An important aspect of data privacy is the identification of potential sources of privacy leakage throughout different stages of data processing in smart industries. The vulnerabilities to data breaches exist from: data collection, storage, and sharing to analysis, outsourcing, and transmission. Their vulnerabilities are exploitable by both internal and external actors making it a multifaceted challenge that necessitates comprehensive safeguards. Moreover, the motivations behind cyberattacks may include industrial espionage, aimed at stealing intellectual property and securing a competitive advantage.

Notably, several high-profile cyber incidents have highlighted the substantial losses that result from these data privacy risks. The infamous NotPetya attack on Merck, which incurred losses exceeding $310 million, serves as a stark example [3].

To address these data privacy concerns and to secure sensitive data the following strategies can be used:

- Data Classification: classifying data based on its sensitivity and confidentiality.
- Data Transformation: Implement techniques like obfuscation, anonymization and encryption.
- Secure Search: Utilizing private information retrieval also known as PIR and secure ranked keyword search to search and retrieve transformed data securely.
- Data Access Control: define and enforce access controls based on user roles and privileges.
- Privacy Management Techniques: The paper presents various techniques to manage privacy in smart manufacturing, including TPM, PawS, Privacy by Design, Data provisioning, etc.
- Addressing privacy concerns in smart factories involves exploring key technologies like the Internet of Things (IoT). Six types of data privacy concerns in IoT applications include:
- Identity Privacy: protecting ownership information to safeguard identities.
- Location Privacy: Ensuring location data privacy with explicit user consent.
- Search Query Privacy: Preventing exposure of personal information through IoT queries.
- Digital Footprint Privacy: Securing digital trails to maintain operational privacy.
- Personal Behavior Privacy: Addressing concerns about data collection without consent.

- Personal Health Data Privacy: Mitigating potential misuse of sensitive health data by health insurers [4].

Privacy and the integrity of data have become of utmost importance in the era of Industry 4.0. With interconnected devices generating vast amounts of data, compliance with data protection regulations such as GDPR and CCPA is crucial. Data breaches have been on the rise, exposing millions of records. To secure data in Industry 4.0, organizations are using practices like robust security measures, data encryption, anonymization techniques, and regular audits. Common data privacy regulations like GDPR, CCPA, and HIPPA each have their specific focus, aiming to safeguard individual data rights [5].

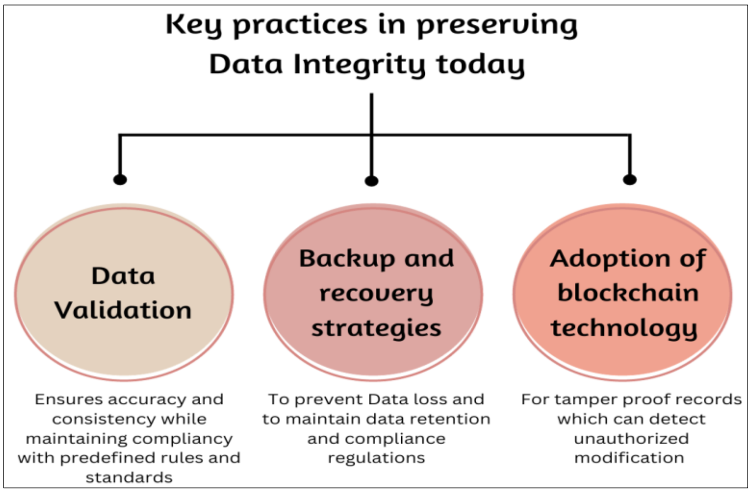
Key practices in preserving Data Integrity today
Data Validation
Backup and recovery strategies
Adoption of blockchain technology
Ensures accuracy and consistency while maintaining compliancy with predefined rules and standards
To prevent Data loss and to maintain data retention and compliance regulations
For tamper proof records which can detect unauthorized modification

Figure 1 Preserving Data integrity [6].

Amid the complexities of Industry 4.0, managing third-party risks, especially when collaborating with data processors, becomes crucial. Especially in contexts like cloud computing and outsourcing. Building a robust data privacy culture is foundational, involving the training of employees, and the establishment of governance frameworks. By making privacy an integral part of the corporate culture, organizations can better adapt to the evolving landscape of Industry 4.0 and meet the requirements of emerging technologies and changing regulations.

1.2. AI-Driven Data Analytics in Industry 4.0

The incorporation of AI-driven data analytics with the developing perspective of Industry 4.0 has paved the way for incomparable efficiency and innovation. Implementing the various data analytic techniques in Industry 4.0 creates an abstruse impact on the entire industry. As Industry 4.0 transpires, artificial intelligence stands as a standing ground that builds organizations to mobilize the extensive amount of data generated by associated systems [7].

Big data analytics are fundamental to AI and machine learning. The real value of the data is only conceived when it is handled and analyzed by AI and machine learning methods, leading to data-driven decision-making [8]. The key applications of AI and machine learning enable it to recognize the analysis of significant data, learn patterns from advanced algorithms, and meet maintenance requirements. These applications and processes play a crucial role in the data analytics present in Industry 4.0 by enhancing quality, reducing costs, improving efficiency, and optimizing inventory management throughout the manufacturing procedure [9].

Protecting the privacy of large volumes of data has been a critical challenge for AI-driven data analytics. Protection of big data in Industry 4.0 is a tedious task but can be resolved by the introduction of techniques such as anonymization, de-identification, data suppression, and differential privacy [10].

Anonymizing data does not guarantee that it is always completely impossible to uncover the identity of the individual. Some of the common data anonymization techniques include: pseudonymization, where data accuracy and integrity are preserved when a private identifier is replaced with the fake identifier. Generalization is when data is made anonymous from removal of certain data. Data masking ensures that data is hidden through encryption, shuffling or substitution. Synthetic enables datasets to be fabricated using pre-modeled algorithms while the original dataset is in use. Perturbation is where sensitive data is hidden by the use of customized noises [11].

Some of the techniques mentioned can be: reversible, irreversible compression, and cross-referencing. With the use of various formulas and theories, the most efficient method can be used by a service to get the best results [10]. De-identification is most commonly used in Industry 4.0 to protect a user's privacy, follow data protection regulations and ensure secure handling of data in various technologies. The specific procedures and techniques employed for de-identifying information can differ based on the situation but must be chosen to effectively safeguard the privacy of individuals [11]. With this, key aspects of de-identification in Industry 4.0 include: data privacy, compliance, data security, data sharing, machine learning and analytics.

Along with the aspects of de-identification, there are methods that can be followed, such as: record suppression, cell suppression, randomization, shuffling, sub-sampling, and aggregation.

Some more methods include: Adding Noise, Character Scrambling, Character Masking, Truncation, Encoding, Blurring, Masking, Perturbation, Redaction and Creating Pseudonyms or Surrogate. With this de-identification can be made successful in the desired workspace [10].

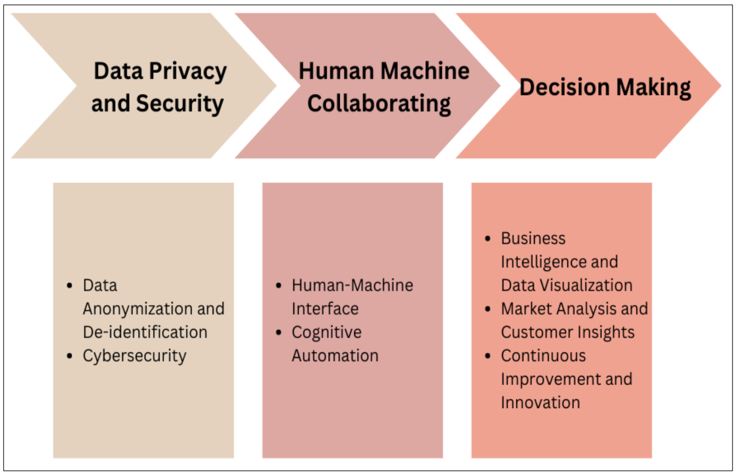
Data Privacy and Security
Human Machine Collaborating
Decision Making
Data Anonymization and De-identification
Cybersecurity
Human-Machine Interface
Cognitive Automation
Business Intelligence and Data Visualization
Market Analysis and Customer Insights
Continuous Improvement and Innovation

Figure 2 Basis of AI-Driven Data Analytics in Industry 4.0 [12].

2. Research Methodology

Adopting the Systematic Literature Review, also known as SLR methodology, in our pursuit of a thorough exploration and synthesis of existing knowledge. This choice is driven by several key considerations.

- Holistic Examination: SLR allows us to systematically review a wide range of literature for a comprehensive assessment
- Rigorous Evaluation: With predefined inclusion and exclusion criteria, SLR ensures a rigorous evaluation of our selected sources, promoting the quality of our findings
- Evidence-Based Analysis: Following the SLR methodology grounds our findings in empirical research, enhancing the credibility of our study
- Gap identification: Through SLR, identifying existing knowledge gaps and areas requiring further investigation, providing valuable insights for future research endeavors.

2.1. Search Strategy

An extensive search conducted between August 2023 and October 2023 was made to find terms related to privacy, data protection, Industry 4.0, and AI. With these terms, an extensive literature review of the topic was managed. A search was coordinated using distinct search terms related to the topic.

2.1.1. Inclusion and Exclusion criteria

To govern the selection of relevant papers, establishing a set of carefully defined inclusion and exclusion criteria will serve as a fundamental framework to ensure systematic and rigorous examination of literature pertinent to our research on privacy and data protection challenges within Industry 4.0.

Inclusion Criteria

- Papers written in English are considered.
- Only peer-reviewed research papers, conference papers and articles are included.
- Each paper must directly address one or more of our research questions
- Papers that focus strongly on data privacy in Industry 4.0, offering insights or solutions
- Papers presenting real-world data or experiments
- Papers that consider different perspectives, like law, ethics, or engineering
- Papers with insights from various industries
- When multiple versions of a paper exist, the most recent version is prioritized.

Exclusion Criteria

- Papers in languages other than English are excluded.
- Papers that primarily offer overviews or surveys of AI in cybersecurity without adding original insights are omitted.
- Papers representing the same work in different conferences or journals are excluded to prevent redundancy.
- Papers that lack sufficient details for comprehensive analysis are excluded.
- Papers promoting or evaluating commercial products or services related to AI and Industry 4.0 without substantial research content are not considered.

3. Analyzing the Key Data Security and Privacy Challenges

As organizations adopt interconnected cyber-physical systems, numerous privacy concerns and cybersecurity challenges arise. Exploring the key concerns in Industry 4.0, including expanded attack surfaces, limited security awareness, reliance on third-party vendors, IoT device vulnerabilities and net attack vectors. Addressing these challenges is crucial for both safeguarding data and maximizing Industry 4.0's transformative potential.

- Increased attack surface: Interconnected systems increase vulnerability [13].
- Lack of security awareness: Industry 4.0's fast tech integration outpaces cybersecurity expertise, posing vulnerabilities. Urgent need for comprehensive cybersecurity understanding to counter cyber threats [14].
- Industry 4.0 relies on third-party vendors, risking security if their practices are weak. Lapses in one vendor's security can have amplified consequences, stressing the need for robust measures across the supply chain [15][16][17].
- Vulnerable IOT Devices: Many IoT devices lacking security measures become prime targets [18].
- New Attack vectors: Challenges demand specialized security measures [19].

While analyzing theoretical information, real-world case studies serve as stark reminders of the potent repercussions stemming from privacy concerns and cybersecurity challenges.

Examining notable instances like the ThyssenKrupp Cyber Attack in 2016. ThyssenKrupp, a German conglomerate specializing in heavy industry, experienced a hacking attack orchestrated by hackers from southeast Asia. The cybercriminals aimed to pilfer "technological know-how and research results from divisions handling orders planning for industrial plants and steelworks in Europe [20].

Similarly, the Merck Ransomware attack of 2017, attributed to the North Korean Lazarus Group, exemplifies the financial risks entailed in the interconnected web of Industry 4.0. As this American pharmaceutical powerhouse grappled with substantial financial losses, the incident illuminated the high-stakes nature of cyber threats and the imperative for robust cybersecurity measures [21].

The Saudi Aramco cyberattack of 2019, attributed to the Iranian hacking group APT33, stands as a harrowing testament to the catastrophic consequences of cybersecurity breaches in the industrial sector. This major oil and gas company experienced billions of dollars in damages and severe disruptions to production, serving as a stark reminder that the industrial landscape is not impervious to the far-reaching impacts of cyber intrusions [22].

In Industry 4.0, case studies highlight the urgency for robust security frameworks against cyber threats. The dichotomy between real-time and non-real-time data presents challenges, with large enterprises adopting dynamic strategies to fortify security in both categories. Let's delve into the dynamic approaches employed by industry leaders to protect these distinct data categories.

Real-time Sensitive Industrial Data and privacy security:

- Predictive Maintenance Algorithms: Leveraging data, and analytics to forecast when equipment is likely to experience issues, allowing organizations to proactively schedule maintenance activities such as M-RAPPOR (Multiple Randomized Aggregatable Privacy-Preserving Ordinal Response) algorithm for enhanced privacy while reducing computational costs [23].
- Privacy-Preserving Techniques: Integration of differential privacy during data collection, utilizing Bloom filters and hash functions for efficient data mapping.
- Zero Trust Architectures: Continuous verification of users, devices and network traffic to prevent lateral movement in case of security breaches [15].
- Encryption for Data in Transit: Implementing robust encryption protocols to secure data during transmission, safeguarding against interception or tampering [17].

Non-Real Time Sensitive Industrial Data Security:

- Cloud-Fog Collaborative Storage: For non-real-time data, typically stored with third-party cloud service providers, a collaborative storage approach is adopted. Data protection schemes combining AES encryption and Reed-Solomen (RS) encoding ensure multilayer security.
- Data Lifecycle Management: Since the data lifecycle is crucial for non-real-time sensitive data. Companies employ Data Loss Prevention (DLP) tools to scan networks for sensitive information. Unauthorized storage locations are identified, and actions like encryption or deletion are taken to maintain compliance with data protection regulations.
- BYOD policies with restrictions: As Bring Your Own Device (BYOD) policies gain popularity for increased flexibility, companies implement strict guidelines for data transfer to personal devices. Device control policies ensure that only secure devices are trusted, aligning the security level of personal devices with company requirements [24].
- Effective Encryption Practices: Extensive use of encryption for various endpoints, including hard drives, USBs, smartphones, and external storage to address remote work challenges [19].

The tailored strategies acknowledge the distinct characteristics of each type of data, ensuring a comprehensive and adaptive approach to industrial data protection. To underscore the importance of robust data protection strategies, analysis of a few notable case studies with severe data breaches has been done

Industry 4.0's data security and privacy challenges require a dual approach: leveraging AI-driven mitigation techniques to proactively identify vulnerabilities and deploying dynamic privacy measures. This proactive stance ensures privacy is integral throughout the data processing lifecycle. Simultaneously, adherence to robust regulatory frameworks like GDPR and CCPA forms a crucial defense. These approaches create a comprehensive strategy to fortify privacy, enhance data security, and build stakeholder confidence in the industrial landscape.

4. First Pillar of Defence: privacy-preserving techniques and strategies from AI standpoint

The first pillar of defense in the dual-pronged approach, focuses on the innovative methods and strategies that are instrumental in ensuring data privacy while maximizing the benefits of AI technologies. It lays the foundation for a responsible and secure AI landscape. For each application in the AI perspective, the different privacy preserving techniques and strategies are listed below in great detail.

4.1. Differential Privacy

Differential privacy safeguards personal data in AI, ensuring privacy across sectors. Ir protects sensitive information during analysis and model training, promoting fairness, transparency, and trust. Compliant with regulations, it mitigates bias and enables collaboration through the sharing of aggregate information without compromising individual privacy [25]. Differential privacy strengthens AI model protection, defending against breaches and adversarial attacks. In research, it facilitates studying privacy-preserving mechanisms, crucial for advancements in AI and privacy, especially in sectors like healthcare and government [26].

4.1.1. Local Differential Privacy

The critical role of artificial intelligence in addressing complex problems in IoT-driven applications, such as image classification, natural language processing, and speech recognition. It highlights the vulnerability of AI and DL models trained on sensitive data, which can lead to privacy inference attacks and privacy leaks. The introduction of a new local differentially private (LDP) algorithm named LATENT, which aims to redesign the training process and maintain high utility compared to existing LDP protocols [27]. The proposed algorithm suits IoT-driven cloud-based environments and demonstrates excellent accuracy with high model quality even under low privacy budgets. There are various methods which go with local differential privacy to level the efficiency of the techniques:

- Laplace Mechanism: The addition of Laplace-distributed noise to each data point.
- Exponential Mechanism: This method perturbs data by sampling from an exponential distribution, providing differential privacy for selecting the output of a function.
- Randomized Response: Users respond to queries or data collection with randomized answers.
- Local Randomization: Locally randomizing data before sharing it.
- With LDP also comes plenty of uses in its appropriate fields, such as:
- Mobile and Edge Computing: Allows data to be processed and shared privately, preserving individual privacy while enabling collaborative AI tasks like federated learning.
- Privacy-Preserving Data Collection: Ensures that user data remains private during data collection.
- User-Centric Privacy: Users engage in AI apps without compromising privacy, ensuring shared data protection and control.
- Location-Based Services: In location-based services, LDP can be used to protect the privacy of users' locations while enabling location-based recommendations and services [28].

4.1.2. Global Differential Privacy

It is a valuable tool in the field of artificial intelligence (AI) as it addresses critical issues such as privacy violations, security concerns, model fairness, and communication overheads. It can be applied in various areas of AI, including machine learning, deep learning, and multi-agent systems, to ensure the privacy of individuals' data and improve the overall performance of AI systems.

In the context of machine learning, global differential privacy mechanisms are used to preserve the privacy of participants in a dataset, ensuring that the aggregate output of a query function does not reveal sensitive information about any individual's data within the dataset. It can also improve the security of AI systems by reducing the impact of malicious participants and ensuring that the learning algorithm is unaffected by modifications to individual records in the training data [25].

With this in mind, some of the major methods are listed below:

- Noise Injection: One common method is adding random noise to data, query responses, or model parameters. Laplace and Gaussian mechanisms are often used for this purpose.
- Secure Multi-Party Computation (MPC): MPC enables joint computation while safeguarding individual data privacy which is crucial for privacy-sensitive scenarios.
- Homomorphic Encryption: This method allows computations on encrypted data without revealing the underlying data, making it useful for privacy-preserving AI.

4.2. Federated Learning

Federated Learning (FL) with enhanced privacy mechanisms to facilitate AI model training and data analysis while preserving individual data privacy, particularly in industrial settings. Federated Learning includes secure aggregation mechanisms that ensure the privacy of local model updates while aggregating them to form a global model [29].

The proposed scheme aims to ensure the privacy of individual participant's information during the entire training process. It utilizes techniques such as differential privacy, Gaussian mechanism, Learning with Error (LWE), and Additively Homomorphic Encryption to ensure the privacy of training data during and after the training process. The scheme involves perturbing local gradients, encrypting the perturbed gradients, and embedding them into A-LWE error terms to achieve secure aggregation.

4.2.1. Horizontal Federated Learning

Horizontal Federated Learning (HFL) applies to scenarios where different datasets share the same features but have distinct sample sets. Security mechanisms, like aggregation schemes and encryption, protect user privacy. Multitask-style HFL allows sites to work on separate tasks while sharing knowledge. Security primarily considers an honest server but faces additional privacy challenges with malicious users [30]. The summarization of horizontal federated learning can be denoted as: $X_i = X_j$, $Y_i = Y_j$, $I_i \neq I_j$, $\forall D_i, D_j, i \neq j$. Methods involved in Horizontal Federated Learning are:

Secure Aggregation: This ensures that user contributions remain confidential.

- Homomorphic Encryption: Secure aggregation, allowing computations on encrypted model parameters without revealing the underlying data.
- Multitask Learning Models: This approach helps address high communication costs, stragglers, and fault tolerance issues in distributed AI.
- Client-Server Structure: Data is partitioned by users, and models built on client devices collaborate at the server to build a global federated model.
- Deep Gradient Compression: This enhances communication in HFL between involved entities.

4.2.2. Vertical Federated Learning

Vertical Federated Learning (VFL) is designed for datasets with the same sample IDs but differing feature spaces, often in scenarios like two companies with overlapping user sets but distinct data. VFL enables the collaborative building of prediction models while preserving data privacy through secure computations. Security in VFL assumes non-colluding parties, with at most one compromised by an adversary. A semi-honest third party (STP) can be introduced, which does not collude with either party. The summarization of horizontal federated learning can be denoted as: $X_i \neq X_j$, $Y_i \neq Y_j$, $I_i = I_j$ $\forall D_i, D_j$, $i \neq j$. To achieve Vertical Federated Learning, the following methods are followed:

- Enhancing Data Privacy: This is essential in AI applications involving user data, finance, and healthcare.
- Cross-Domain Learning: Common predictive model built while maintaining the confidentiality of their data.
- Optimized AI Models: VFL leverages the collective knowledge and feature sets of different entities to build AI models with improved accuracy, even when the datasets are vertically partitioned [31].

4.3. Homomorphic Encryption

The fusion of Homomorphic Encryption and Artificial Intelligence presents a promising intersection of cutting-edge technologies. HE, a cryptographic technique that allows computations to be performed on encrypted data without revealing the underlying information, brings a new level of data privacy and security to AI applications. This amalgamation, however, is not without its challenges and intricate solutions. AI Models Connected with Homomorphic Encryption:

- Secure Inference: HE enables secure inference in AI models.
- Private Machine Learning: AI models, such as decision trees or neural networks, can be trained.
- Federated Learning: HE plays a role in federated learning by allowing models to be trained collaboratively on encrypted data from multiple sources.
- Secure Deep Learning: Homomorphic Encryption (HE) enables secure, private training of deep learning models for image analysis, speech recognition and natural language processing [32].

The integration of Homomorphic Encryption and Artificial Intelligence presents a transformative paradigm where advanced AI capabilities are harnessed while ensuring the highest levels of data privacy and security. Despite the challenges, innovative solutions are emerging, making this fusion a frontier for secure and privacy-conscious AI applications [33].

4.3.1. Partially Homomorphic Encryption

The intersection of Partially Homomorphic Encryption (PHE) and Artificial Intelligence (AI) represents a remarkable convergence of advanced cryptography and machine learning. PHE, a cryptographic technique that allows specific mathematical operations on encrypted data, serves as a powerful tool for enhancing the privacy and security of AI applications [34].

In the context of privacy-preserving techniques for machine learning, there are two primary approaches. The first involves using data anonymization techniques to generalize and suppress characteristic features in the dataset. After anonymization, the data can be safely shared with other parties for data analytics. The second approach employs cryptographically secure multi-party computation algorithms to perform computations on encrypted data, ensuring that the same results are obtained as if working with the plain data. The challenges in classical cryptographic methods include key management and the potential risk of the cloud server having access to private data [33].

4.3.2. Fully Homomorphic Encryption

Fully Homomorphic Encryption (FHE) has gained recognition as a groundbreaking cryptographic approach, often deemed the "holy grail of cryptography." It offers a revolutionary solution to cybersecurity challenges by allowing a third party to process encrypted data without accessing confidential information. In the context of AI and security against the potential quantum computer threat, FHE serves as a promising tool. It finds significant utility in privacy-preserving machine learning models, ensuring both data security and accurate analysis. Bootstrapping refreshes noisy ciphertext by evaluating the decryption function homomorphically. FHE schemes are considered bootstrappable if they can evaluate their own decryption algorithm circuit, enhancing their security and practicality [35].

4.4. Secure Multi-Party Computation

Secure Multi-Party Computation (SMPC) is a cryptographic paradigm that allows distributed computation among multiple parties while maintaining the privacy and security of their inputs. It finds application in various tasks, both cryptographic and non-cryptographic, such as machine learning, authentication, electronic voting, and contract signing. The core security requirements for SMPC protocols include privacy, correctness, independence of input, guarantee of output, and fairness, ensuring that participants can collaborate securely even in the presence of dishonest or corrupted entities [36]. SMPC, operating under security models like Semi-honest, Malicious, and Covert Adversary, ensures privacy in collaborative computing. Techniques such as Fully Homomorphic Encryption and Function Secret Sharing offer varied trade-offs in complexity and efficiency. The SyMPC library, based on Function Secret Sharing, exemplifies SMPC's practical use, particularly in machine learning inference.

4.4.1. Secure Two-Party Computation

STPC, a subset of SMPC, involves two parties jointly computing functions over their private data while maintaining privacy. In AI, STPC is essential for secure collaborative model training and analysis. Methods like garbled circuits, oblivious transfer, and homomorphic encryption enable privacy-preserving AI applications, such as confidential financial modeling, secure federated learning, and privacy-centric predictive analytics. Some of the features in connection with AI includes:

- Communication Efficiency: The protocol achieves remarkable efficiency by requiring just 2 ring elements per multiplication in the online phase.
- Round Complexity Reduction: The protocol minimizes the number of online rounds needed for various operations, which enhances the efficiency and speed of the computations.
- Matrix Multiplication Support: Efficiently supported by extending the STPC multiplication protocol to accommodate vector operations, streamlining this critical operation.
- Depth-Optimized Circuits: The protocol optimizes depth in circuits, improving efficiency and reducing online communication, this is done by combination of various AND gates [37].

4.5. Data Anonymization

Data anonymization for artificial intelligence involves the process of transforming or masking sensitive and personally identifiable information (PII) within a dataset to protect individual privacy while maintaining the utility of the data for AI and machine learning purposes. This is crucial for complying with data protection regulations and addressing privacy concerns.

The process typically includes techniques such as generalization, perturbation, and data masking to de-identify information. However, achieving a balance between privacy and data utility is essential. Machine learning models require quality data to make accurate predictions and analyses. Therefore, data anonymization methods should minimize information loss while ensuring that individual data subjects cannot be re-identified [38].

Some of the techniques which are commonly used are:

- Data Masking: Replaces sensitive information with fictitious data, maintaining data format while safeguarding sensitive data for analysis or testing.
- Pseudonymization: Replaces identifying fields with pseudonyms, making it harder to identify individuals but still useful for research and analysis.
- Synthetic Data: Artificially generated data closely mimicking real-world data for safe use in data analysis, machine learning, and testing.
- Data Perturbation: Adds random noise or errors to data to protect privacy without compromising data utility.
- Data Swapping: Exchanges data between individuals or entities while maintaining statistical properties, enhancing privacy for research and analysis [39].

4.5.1. Tokenization

Data Tokenization is an advanced security method that replaces sensitive data with untraceable tokens, ensuring robust data protection without compromising usability. As AI evolves, traditional security measures fall short. Tailored for AI, tokenization coupled with threat detection harnesses the potential of artificial intelligence to bolster data security and identify risks effectively. This synergy marks a cutting-edge approach, striking the perfect balance between data privacy and AI-driven insights [40].

Tokenization which is AI-driven and intrusion detection go in hand with these following points:

- Dynamic Tokenization: AI enables real-time token generation based on context and risk assessment, bolstering security.
- Threat and Anomaly Detection: AI monitors tokenized data and networks, detecting anomalies and potential threats, aiding breach prevention.
- Behavioral Analysis: AI, combined with behavioral analysis, identifies suspicious user behaviors and access patterns using legitimate tokens.
- Real-time Threat Response: AI immediately responds to security issues, minimizing impact.
- Continuous Learning: AI adapts tokenization and threat detection processes by learning new threat patterns.
- Integration with Security Ecosystems: AI seamlessly integrates with existing security systems and tools, enhancing overall protection [41].

4.5.2. Data Masking

Data masking is a critical practice in the modern data landscape, driven by the increasing need for data privacy and regulatory compliance. Data masking ensures that sensitive data is concealed or altered in a way that it cannot be easily traced back to individuals. It helps organizations maintain data privacy control mechanisms, reducing the risk of personal information leakage. By using techniques like shuffling, encryption, and masking, data masking safeguards sensitive information while allowing businesses to use the data for various purposes, including analytics and research. In a world where data-driven decision-making and artificial intelligence are on the rise, data masking is crucial for striking a balance between data utility and privacy protection [42].

4.6. Privacy Preserving Machine learning

Privacy Preserving machine learning, a crucial domain in AI, addresses data privacy concerns. It employs innovative strategies like differential privacy, federated learning, secure multi-party computation, and homomorphic encryption. These methods safeguard data confidentiality while enabling model development and deployment. Crucial in healthcare, finance, and other sectors, they ensure data security and compliance with privacy regulations. Based on robust cryptographic and statistical principles, they prevent unauthorized data exposure [43].

4.6.1. Privacy Preserving Neural Networks

PPNN in deep learning and machine learning addresses securing and preserving data privacy in collaborative prediction settings involving multiple data providers. Unlike traditional models assuming single-source data, real-world scenarios involve contributions from multiple providers, each with partial features of a complete sample. This enables cooperation without exposing private information. To tackle this, PPNN models employ multi-client inner-providers to encrypt and transmit data to a trained model on a remote cloud server. The server performs data predictions, ensuring data privacy and security. This approach offers both privacy and security for the data, even when employing different neural network architectures with non-linear activation functions on the remote server [43]. Along with privacy preservation and secure collaboration, the key benefits of PPNN models are in many of the proposed models under this umbrella, like the proposed PPNNP model:

- Flexible Neural Network Architectures: PPNN models enable flexibility with diverse neural network architectures, accommodating various real-world scenarios.
- Reduced Computational Cost: PPNN models reduce computational costs compared to traditional methods, especially beneficial for resource-intensive machine learning tasks.
- Accuracy: PPNN models can achieve high prediction accuracy, often exceeding 90% even when using different network architectures and non-linear activation functions.

Reduced Communication Overhead: PPNN models can minimize communication overhead in the protocol, making them suitable for applications with multiple data providers [43]. Computational methods enable neural network computations on encrypted data, primarily through Homomorphic Encryption (HE) and Functional Encryption (FE). Table 1.1 - Encryption Techniques in Privacy Preserving Neural Networks.

Privacy Preserving neural networks, bridging AI and cryptography, secure sensitive data in machine learning. Methods like CryptoNets, CryptoDL, Tanaka's scheme, and pixel-based encryption provide strong foundations. Future advancements aim to improve privacy and performance for secure AI applications in sensitive domains.

Table 1 Methods found from the combination of AI and Cryptography

Model	Introduction	Encryption Technique Used	Key Challenges
CryptoNet	2016	Homomorphic, Encryption	Computational complexity, increased latency
CryptoNN	2019	Function Encryption	Longer training times due to frequent communication for key generation
CryptoDL	2017	Polynomial approximations	Limitations with complex datasets (e.g., CIFAR-10)

4.6.2. Split Learning

Split learning efficiently partitions neural networks horizontally, distributing computations across multiple devices to handle specific segments, effectively overcoming resource constraints.

Horizontal Partitioning: In split learning, the neural network is divided horizontally, effectively slicing it into segments. This divides the neural network into segments for individual device processing, optimizing resource usage and alleviating memory and processing limitations.

Synchronization of Learning Data: Ensures accurate processing across devices by managing data transfer and consistency, employing techniques like model updates and caching for synchronization, even under unreliable network conditions. Split learning offers a range of benefits, making it a powerful tool for resource-constrained devices:

Resource Efficiency: Enables resource-intensive deep neural networks to operate on mobile devices without memory congestion or overloading processors.

Real-time Processing: Facilitates real-time AI processing on mobile devices, maintaining acceptable latency for tasks like object recognition and speech processing [45].

Split learning empowers resource-constrained devices to perform complex AI tasks. By horizontally partitioning neural networks and carefully managing data synchronization, split learning eliminates the resource barriers that previously limited the deployment of deep learning models.

4.7. Model Aggregation

Model Aggregation merges local machine learning models or updates from various sources into a global model, ensuring data privacy. It's vital in privacy-sensitive domains like medical research or financial analysis. Methods like federated learning and secure aggregation enhance model accuracy and generalization, enabling collaborative learning across organizations while safeguarding data privacy, thereby fostering scalable and secure machine learning solutions.

4.7.1. Secure Aggregation of Model Updates

In the context of federated learning, Secure Aggregation plays a pivotal role by enabling multiple participants to collaborate on training machine learning models while safeguarding the privacy of their respective datasets. It achieves this by securely aggregating model updates without revealing individual data specifiers, rigorously protecting the privacy and confidentiality of each participant's data [46]. Secure aggregation relies on secure multiparty computation (MPC) to compute model parameter updates' um securely, preserving user privacy without revealing individual data. Research in secure aggregation encompasses various methods like generic secure MPC protocols, DC-nets, homomorphic encryption, and pairwise masking [47]. In mobile environments with communication dropouts, the protocol minimizes overhead, ensuring robustness and defending against adversarial threats. It employs high-level secure data aggregation protocols using cryptographic techniques, enabling private model training and secure model updates while allowing customization for specific use case requirements.

4.8. Adversarial Attacks and Defense

Adversarial attacks in machine learning (ML) exploit inherent vulnerabilities in machine learning models, using sophisticated techniques like white-box, block-box, and transfer attacks. These methods pose a substantial threat to critical applications like image recognition and autonomous vehicles due to subtle alterations known as adversarial examples.

Assessing the success of adversarial attacks relies on metrics like accuracy and recall, which often prove inadequate in the adversarial context. To combat these attacks, robust defense strategies are essential. These encompass adversarial training, input preprocessing and the development of resilient model architectures. Moreover, the emerging field of Explainable AI (XAI) assumes significance by enhancing model interpretability, a crucial element in effective adversarial defense.

4.8.1. Adversarial Training

Adversarial Training fortifies machine learning models against manipulative inputs by incorporating adversarial examples into training datasets. Techniques like FGSM and PGD iteratively expose models to both normal and perturbed samples, enhancing their resilience and adaptability. This training method optimizes models for accuracy on clean data and resilience against adversarial perturbations. Crafted to deceive models while maintaining perceptual constraints, adversarial examples play a crucial role in securing models. This approach is highly effective in domains like image & natural language processing.

Despite its efficacy, It introduces challenges such as elevated computational demands and potential overfitting to specific attack strategies. Addressing these concerns involves employing diverse perturbations, and ensemble techniques, and optimizing for robustness. Adversarial Training emerges as a sophisticated approach, reflecting the nuanced interplay of adversarial examples and regularization techniques, indispensable for fortifying machine learning models against evolving threats [48].

4.8.2. Robust Machine Learning

In Robust Machine Learning, the primary goal is to enhance model resilience against adversarial perturbations and uncertainties in input data. It involves strengthening models to sustain optimal performance despite distributional shifts. Strategies like adversarial training, regularization, and ensemble techniques are integral for fortifying robustness against adversarial interventions [49].

Robust ML includes outlier identification, data preprocessing, and uncertainty modeling beyond adversarial context. Adversarial training enhances resilience by incorporating adversarial instances into training data using methods like FGSM or PGD. Regularization mechanisms like L1 and L2 help control model complexity and prevent overfitting [50][51]. Ensemble methods, integrating diverse models or perturbation strategies, significantly contribute to robustness by fostering model diversity. Robust ML extends beyond supervised learning, addressing challenges in semi-supervised and unsupervised scenarios where labeled data is limited [49]. Challenges persist in maintaining robustness while optimizing computational efficiency. Research endeavors focus on innovative optimization, transfer learning, and versatile integration of robust ML across domains for adaptable model fortification in dynamic environments [51][52].

4.9. Privacy Preserving AI Governance

This refers to the set of policies, practices and mechanisms that organizations and institutions put in place to ensure that AI systems and processes comply with privacy regulations and protect sensitive data. This governance framework is essential to maintain the trust of users and stakeholders in AI applications.

4.9.1. Policy Based Privacy Controls

Policy-Based Privacy controls oversee rules and guidelines dictating how AI systems manage and safeguard user data, aiming to minimize data breach risks, unauthorized access, and privacy violations.

- Data Handling policies: Govern data collection, storage, processing, and sharing, specifying encryption, access, and retention.
- Consent Management: to ensure compliance with privacy regulations like GDPR and CCPA.
- Data Anonymization & Pseudonymization: To obscure data while preserving utility for analytics.
- Compliance auditing, monitoring and PIAs: Conduct regular audits, PIAs, and monitoring to ensure policy adherence [53].

4.10. Privacy-Preserving Data Sharing

Data sharing in Industry 4.0 encounters privacy challenges due to legal and regulatory hurdles, large-scale consent management, and trade-offs between anonymization and data quality aiming to enable analysis across databases without sharing individual-level data. Frameworks, like the medical domain's Systematization framework, categorize and assess data sharing methods based on various axes. The privacy protection encompasses three key aspects: safe data inputs by anonymizing, aggregating, or encrypting data; secure processing through safe settings; and safeguarding analysis results using techniques like differential privacy. Additionally, the approaches exhibit usefulness like deduplication, reconcile data across databases, flexibility adapts to diverse analysis needs and scalability maintains performance with growing complexity, addressing various data processing challenges.

Categories of Approaches for medical domain: Data sharing approaches have been categorized:

- Distributed analysis: data is exchanged in aggregated or anonymized forms, emphasizing privacy protection. Example include SHRINE/i2b2, DataSHIELD, OMOP/OHDSI, etc
- Cryptographic Secure Multi Party Computation: Ensures high privacy with encrypted data exchange (e.g., MedCo, Sharemind MPC).
- Data Enclaves: Securely store pooled individual-level data for research purposes [54].

Secure data sharing in group communication requires transmitting multimedia data to specific users while safeguarding privacy. While end-to-end encryption secures individual messages, "identity-based broadcast encryption" ensures secure data transmission to authorized recipients while preserving both message confidentiality and user identities. Introducing an "anonymous revocable identity-based broadcast cryptosystem (ARIBBE)" addresses these privacy concerns in data sharing [55]. ARIBBE scheme ensures subscriber privacy, efficient user revocation, and fine-grained data access control evolution. It also enables third-party involvement in the revocation process, aiding in privacy-preserving data sharing.

4.10.1. Secure Data Enclaves

Secure enclaves, or trusted execution environments (TEEs), play a pivotal role in fortifying cybersecurity in Industry 4.0, particularly in safeguarding data and upholding the integrity of AI driven processes:

- AI Models Protection: Secure enclaves shield AI models from tampering and theft.
- Defense Against AI-Attacks: Enclaves use real-time anomaly detection against AI threats.
- Confidential collaborations: enclaves facilitate secure data sharing and AI model training.
- Secure enclaves complement AI in Industry 4.0 by enhancing threat detection and response:
- Behavioral Analytics: AI algorithms within enclaves can continuously analyze user and system behaviors, quickly identifying anomalies that may indicate a cyberattack.
- Predictive Security: AI models running in secure enclaves can predict potential security breaches by analyzing historical data and trends, enabling proactive cybersecurity measures [56].

4.10.2. Data Trusts

Data trusts facilitate secure and equitable data exchanges among stakeholders, acting as intermediaries in a regulated, transparent environment. Key components of Data Trust Include:

- Data Subjects: Entities generating or subjects of the data, which may include personal information or individual content.

- Data Collectors: Institutions or entities that interact with data subjects to collect data, like employers, marketing firms, or social media services.
- Data Users: Third parties, researchers, or agencies requiring data access for analysis, research, or decision-making, following specific trust protocols.
- Data Controllers: Responsible for managing the data trust framework. Administrative controllers formulate policies, while technical controllers handle data visualization [57].

Data trusts, often operated through cloud-based platforms like DTaaS, enable transparent and secure interactions among stakeholders. Data subjects submit data to the trust, while data collectors, interacting with subjects, register on the platform. Data users, intended data recipients, create or have profiles created on the trust. The trust's data controller layer securely stores the data, managing data access even if subjects don't directly provide it to collectors [58].

Data trusts apply to diverse scenarios like COVID-19 vaccination or research data sharing, providing insight into data use by utility and water providers. Research goals involve designing transparent multi-partner data sharing, integrating blockchain for accountability, and evaluating policy effectiveness. Data trusts offer forward-thinking solutions for complex data sharing and privacy challenges. Preserving privacy in AI-driven data analytics is crucial. Techniques secure sensitive data during analysis by keeping it confidential [59].

5. Second Pillar of Defense: Regulatory Frameworks and Compliance

Delving into the second pillar of data protection, the focus is shifted from the practical aspects of securing sensitive industrial data to the legal and regulatory framework that governs the crucial aspect of modern industry. Data breaches and privacy violations affect finances and tarnish company reputations, eroding client trust. Following are the explored regulatory frameworks and compliance practices:

5.1. General Data Protection Regulation (GDPR)

The EU General Data Protection Regulation (GDPR), one of the most stringent privacy laws, regulates the processing and transfer of personal data of individuals within the EU. It applies not only to EU-based entities but also to organizations outside the EU handling the data of EU residents. Obtaining valid consent from individuals prior to processing their data is a crucial requirement for GDPR compliance.

- Consent and Transparency: Clear consent and transparent data usage empower informed choices.
- Data Protection Impact Assessments (DPIAs): Identifies privacy issues during data processing.
- Data Subject Rights: Individuals gain control to access, rectify, or erase their data.
- Data Portability: Enables receiving personal data in machine-readable format.
- Data Breach Notification: Swiftly notifies authorities and individuals about breaches [4].

GDPR grants favorable exceptions for research, including scientific and public health studies, provided appropriate safeguards are implemented, allowing flexibility in data protection principles and data subject rights. Expectations to data protection principles include:

- Purpose Limitation: simplifies research by permitting personal data reuse without new legal basis.
- Storage Limitation: Research data can be kept longer for result verification beyond initial purpose
- Special Categories of Data: Sensitive data in research needs lawful authorization, proportional to aims [60].

For the processing of special categories of sensitive data, researchers can rely on Article 9(2)(j) to provide an exception to general prohibition, although it should be applied cumulatively with Article 6 to ensure comprehensive data protection. GDPR shapes AI by restricting EU personal data use, ensuring access to automated decisions, safeguarding sensitive research data, addressing bias, and facing challenges in evolving smart cities and AI systems [61].

5.1.1. California Consumer Privacy Act (CCPA)

The CCPA in California grants consumers the right to access their personal data held by companies and request details on third-party data sharing. It aligns with GDPR objectives but introduces its own set of privacy requirements.

- Right to Know & Deletion: Control over data enables awareness & removal if no longer needed.
- Data Sale Opt-Out: Enables individuals to prevent unauthorized data monetization.
- Data Minimization: Encouraging organizations for limited data collection for specific purposes.

- Non-Discrimination: forbids discriminating against privacy-right exercising consumers [62].

Data minimization in AI development, crucial for privacy, requires collecting only essential data, potentially affecting model quality. CCPA's access and deletion rights may lead to reliance on older data, balancing privacy compliance with effective model training [63]. Impact on targeted advertising due to CCPA leads AI algorithms relying on extensive consumer data to adapt. Increased compliance costs can be an issue which can be solved using AI tools Third-party vendors and data providers aid AI model development. Transparency and accountability are vital for user data use in AI development [64].

5.1.2. Information Technology (IT) Act, India

In India, the Information Technology (IT) Act is the main framework that establishes provisions pertaining to data breaches since there is no separate legislation for collecting, processing, utilizing, and storing data by organizations. It is based on the United Nations Model Law on Electronic Commerce. Businesses that harness AI's potential within the framework of Industry 4.0 will be better positioned to prosper in the rapidly changing modern industrial environment.

- Data Protection and Security Obligations: Best security practices and procedures to save data.
- Sensitive Personal Data and Information (SPDI): Special care handling of sensitive data.
- Data Breach Notification: Smooth responses to data breach and mitigation potential harm.
- Extraterritorial Jurisdiction: Data breaches attacked towards Indian citizens can lead the Act to hold jurisdiction.
- Penalties for Non-Compliance: Penalties for non-compliance, including financial penalties and potential imprisonment [65].

Data protection and privacy rely on the Personal Data Protection Bill (PDPB), legislating these aspects. Ensuring consent and transparency, especially with sensitive user information is crucial. AI systems handling such data must meet stringent security requirements to prevent breaches and cyberattacks. Addressing responsibility for AI errors aligns with legal considerations under the IT act, which includes provisions on intellectual property rights impacting AI, notably patents, copyrights, and software.[66] Mandated by data protection regulations, MultiFactor Authorization (MFA) enhances security against identity theft and financial loss from potential breaches. Despite its effectiveness, implementing MFA encounters numerous challenges.

IEEE and IEC, prominent global organizations, establish standards fostering seamless technology integration across various sectors—energy, telecom, healthcare, and IT—ensuring interoperability, safety, and efficiency. IEEE significantly influences electrical and electronic engineering standards, impacting data privacy and security. Simultaneously, IEC, a global body in electrotechnology, shapes international standards with a significant impact on data privacy and security as well.

- IEEE P2830 - Recommended Practice to secure sensitive data in Medical Devices: Guidelines secure sensitive medical device data, crucial in healthcare's priority of patient data security [67].
- IEEE 802.1X- Port Based Network Access Control: Widely used for secure network access, allowing only authorized users and devices, fundamental for data security [68].
- IEEE 802.11i- Wireless Network Security: Enhances Wi-Fi security with advanced encryption, safeguarding wireless data transmission [69].
- IEEE 1680 assesses imaging equipment environmentally, while IEEE 1547 connects power systems; both bolster data security and privacy.
- IEC 62443-Industrial Communication Networks-Network and System Security: Secures industrial automation, vital for protecting sensitive data in manufacturing and critical infrastructure [70].
- IEC 61701 - Protection of Information in IEC Open Systems: Standard focuses on data protection in open systems, stressing security for safeguarding sensitive data across diverse platforms [71].
- IEC 80001-Application of Risk Management for IT Networks for Medical Devices: In healthcare, these standard stresses risk management for securing patient data privacy in medical devices [72].
- IEC 27001 - Information Security Management System (ISMS): IEC 27001, though not exclusive, is widely recognized for its holistic approach to information security. It establishes, implements, and maintains an ISMS, safeguarding sensitive data [73].
- IEC 62645 focuses on data structures in process equipment catalogs, indirectly impacting data security in industrial processes. Meanwhile, IEC 61850 prioritizes secure communication protocols in substations, crucial to maintain data integrity and protect power infrastructure [74][75].

Transparency is vital in AI systems to prevent biased algorithms that perpetuate discrimination, impacting both ethics and society. Individuals must have mechanisms to question AI decisions and understand their underlying rationale for ensuring fair and accountable AI implementation.

6. Limitations in current Regulatory Framework in Industry 4.0

Existing regulations, like India's IT Act of 2020, may not be agile enough to address emerging technology, highlighting the need for continuous updates to balance blocking dangerous online content with preserving free expression and consumer privacy as Internet usage in India increases [76]. Each technology pillar within industry 4.0 presents distinct legal challenges.

- Cloud Computing: Raises concerns regarding jurisdictional boundaries and privacy violations.
- Internet of Things: Stresses governance, policy, security for device interconnection management.
- Big Data Analytics: Spotlights challenges in meeting data protection, IP rights & competition laws.
- Cyber-Physical Systems: Advocates a specific legal framework for autonomous systems [77].

Global data protection regulations like GDPR and CCPA impact multinational companies handling consumer data across jurisdictions, complicating cross-border data transfers in Industry 4.0. The use of AI in processing extensive data raises privacy concerns, demanding attention to regulatory gaps arising from Industry 4.0's concepts not aligning with existing frameworks. Several contributing factors exacerbate policy challenges like limited understanding of emerging tech, policy uncertainties for policymakers, differing impacts of disruptions across societal sectors leading to stakeholder conflicts and regulators navigating the fine line between overregulation and under-regulation to protect interests [78]. There should be a departure from traditional regulatory mechanisms toward a more adaptive regulatory approach. Highlighting conventional regulations' inadequacies with evolving tech and proposing a shift:

- Adaptive Regulation: Responsive and iterative approaches to regulation.
- Regulatory Sandbox: Creating environments to prototype and test new approaches.
- Outcome-based Regulation: Focusing on results rather than rigid forms.
- Risk-weight Regulation: Tailoring regulation based on data-driven assessments.
- Collaborative Regulation: Engaging a broader set of stakeholders for national and international alignment [79].

In most nations, governments supply the infrastructure for the digital world. However, there's a lack of clarity on Industry 4.0 which hinders infrastructure transformation roadmap. Malaysia's National Policy (Industry4WRD) is a critical step towards becoming a strategic partner in smart manufacturing and related services in the Asia Pacific. The introduction of Regulatory Sandbox initiatives enabling firms to manage regulatory risks during the testing stage and fostering innovation [79][80]. Local industries, especially SMEs, struggle adopting Industry 4.0 due to awareness gaps, low digital adoption, talent shortages, evolving customer behavior, cost misconceptions, and few local success stories with such tech [81].

7. Field research

Digital ecosystem of Industry 4.0 presents significant privacy risks due to extensive data generation and interconnected systems. This section examines real-world cases, including many multinational companies

7.1. Data Overload

A foundational challenge in Industry 4.0 is the voluminous data generated from various sources. This data is invaluable for enhancing operations and services but prompts concerns about the extent of data collection and its potential privacy implications [6]. Examples of organization who struggled

Cases: Marriott International Data Breach & British Airways Data Breach

The 2018 cyberattack on Marriott International, comprising data for over 500 million guests, exposed critical security lapses, including a prolonged undetected breach dating back to 2014 within the Starood network. The incident resulted in a £99 million GDPR fine, class-action lawsuits, and projected losses in the billions for Marriott [82]. Similarly, the 2018 British Airways data breach, initially fined $238 million and later reduced to $26 million for GDPR non-compliance, exposed vulnerabilities in the airline's network through a compromised third-party supplier. The attack, lasting six weeks, highlighted the significance of collaborative efforts with partners to mitigate organizational data [83].

7.2. Mitigation strategies

Data minimization involves collecting and retaining only the data necessary for a specific purpose, reducing the associated privacy risks [84][85].

Privacy by Design: A proactive approach, embedding privacy into system development from the start. This ensures that privacy is a fundamental aspect of any project from the outset [86][87].

7.3. Data Sharing with Third Parties

Industry 4.0 necessitates data sharing with suppliers, customers, and cloud service providers. While this is crucial for collaboration, it introduces new risks of data breaches and unauthorized access [88].

7.4. Cases: Volkswagen Group & BMW Group

The Audi data breach affecting 163,000 individuals revealed security lapses in sales and marketing data collected from 2014 to 2019, left exposed online from August 2019 to May 2021. This incident underscores privacy challenges in Industry 4.0, simplified by Volkswagen Group's data collection from vehicles and factories, emphasizing concerns about sharing data with undisclosed third parties [89]. BMW Group faced cybersecurity incidents in 2022, with breaches on BMW France's social media accounts and a claim by KelvinSecurity Team accessing and offering for sale data of 384,319 UK BMW car owners. This emphasized the need for data minimization strategies to mitigate privacy risks [90].

7.5. Mitigation Strategies

Employee Education: Training programs like phishing and social engineering are essential.[91].

Collaboration with Industry Partners: Organizations can collaborate with industry peers to develop and implement best practices for privacy and data protection, increasing standards [92].

7.6. Cybersecurity Threats

The increasing targeting of Industry 4.0 systems by cybercriminals poses a significant threat to privacy & data protection. These attacks can lead to data theft, operational disruptions and ransom demands [93].

7.6.1. Case: Siemens Cyber Attack

Siemens, a leading industrial technology company, faces challenges in protecting extensive data collected from its industrial products and systems. A multi-faceted approach, including multi-factor authentication and data encryption, is vital to fortify data security [94][95].

7.7. Mitigation Strategies

Data Encryption: Data encryption safeguards data from unauthorized access by scrambling it so that only authorized individuals can decipher it [96][97].

Multi Factor Authentication: Multi-factor authentication enhances security, demanding multiple factors for authentication, safeguarding systems and accounts against unauthorized entry [95].

7.7.1. Transparency, Accountability, and Enforcement

Transparency in data collection and usage, accountability for privacy violations, and the enforcement of privacy laws and regulations are challenging in the rapidly evolving landscape of Industry 4.0 [98].

Case: The SolarWinds Cyber attack

The SolarWinds cyberattack, attributed to the Russian Foreign Intelligence Service, began in September 2019, compromising SolarWinds' Orion software updates. With nearly 18,000 customers affected, including federal agencies, the breach allowed remote exploitation of compromised systems, highlighting a significant and sophisticated cyber security threat [99].

7.7.2. Mitigation Strategies

Government Regulation: Governmental role is crucial in crafting and enforcing adaptable privacy laws, like the EU's GDPR, in response to evolving tech [100].

The field research conducted on privacy and data protection challenges in Industry 4.0 sheds light on the intricate landscape of threats and solutions within the digital transformation era. Integrating these methodologies into organizational practices and embracing evolving technological solutions are imperative to mitigate risks and foster a secure, ethical, and progressive Industry 4.0 landscape.

8. Cross-Disciplinary Approaches

It is becoming more and more clear that a multidisciplinary, cooperative approach is necessary in the endeavor to handle the complex privacy issues that arise at the intersection of industry 4.0 and AI [101]. The need for interdisciplinary collaboration to address the complex issues arising from the integration of artificial intelligence (AI) and digital technologies in various industries [101]. Addressing the intersection of IP law and antitrust enforcement in AI, the challenge lies in balancing the protection of intellectual property rights with fostering innovation without stifling industry progress [102]. Growing AI adoption underscores the global trend of stricter data privacy regulations like GDPR and CCPA, imposing obligations on companies, including AI users, for data minimization and consent adherence [101]. In Industry 4.0, the shift from traditional reliance on OT and IT to integrating AI in manufacturing introduces cybersecurity risks, emphasizing the need to address and mitigate potential threats to secure smart factories [103]. The review underscores the role of ethical considerations, emphasizing responsible autonomy, transparency, and accountability in AI development, including the OECD's AI principles, guide responsible AI development by addressing ethical challenges [104]. Psychologists and engineers collaborate to integrate psychological insights into AI design, especially in social media, where the algorithm can shape up a user's behavior and information consumption [105]. Collaboration that ensures AI-driven healthcare solutions adhere to ethical, privacy, and regulatory standards, particularly in handling sensitive patient data. This approach helps in the secure integration of AI in healthcare, providing benefits to patients while safeguarding their data and privacy [106]. Underscoring the need for cross-disciplinary collaboration and a comprehensive approach to address privacy and data protection challenges in the Industry 4.0 era are crucial for our ever-evolving world. It emphasizes IP law's intersection with antitrust enforcement, AI's effect on data privacy, the shift in manufacturing security and ethical concerns in AI deployment [101]. Furthermore, the merging of AI and environmental sustainability is a rising concern. As Industry 4.0 technologies, including AI, become more prevalent, their energy consumption and environmental impact need consideration. Collaborative efforts between AI researchers, environmental scientists, and policymakers is vital to create eco-friendly AI solutions and sustainable tech guidelines [107].

9. Future Trends and Scope

As we have explored the current landscape of privacy concerns in Industry 4.0 from an AI driven perspective, it is crucial to pivot our attention toward the future. Industry 5.0 shifts focus to people as the central element in the production sector, aiming for a harmonious blend of technological advancement and human collaboration. This involves utilizing robots and collaborative robots for repetitive tasks, with humans specializing in intellectual production, requiring qualification for proactive participation in the societal model [101].

Swarm Robotics in Manufacturing: A swarm of robots is the coordination of multiple robots that can perform a collective task and solve a problem more efficiently than a single robot. Although research has been going on for several decades, a breakthrough of swarm robotics, especially for industrial applications, has not yet occurred [108]. In the future, it will become a new and powerful tool in precision medicine, allowing personalized therapies such as minimally invasive surgery or direct polytherapy delivery to malignant cells inside the human body [109].

Evolution of Explainable AI: Explainable artificial intelligence (XAI) is a set of processes and methods that allows human users to comprehend and trust the results and output created by machine learning algorithms. This is essential for gaining user trust, meeting regulatory requirements, and ensuring accountability in critical decision-making processes. Explainable AI enables brands to effectively detect flaws in the data model and biases in the data itself [110].

AI-Enhanced Human Augmentation: Augmented Intelligence denotes technologies of AI that do not replace human decision-making, but rather provide additional information for decision-support [111]. AI-enhanced human augmentation integrates AI with wearable devices, prosthetics, and cognitive technologies, optimizing physical and cognitive abilities.

Simultaneous Localization and Mapping (SLAM) Algorithms: SLAM refers to a set of methods to solve the pose estimation and 3D reconstruction problem simultaneously while a system is moving through the environment [112]. Developing SLAM algorithms can enable more precise and instantaneous mapping in dynamic industrial situations.AI can only function to a certain extent due to their set amount of computational power. With the advancements in quantum computing, there is a possibility of significantly enhancing the performance of machine learning and AI [113]. Although it is still in its early stages, quantum computing has the potential to resolve challenging issues that are beyond the reach of conventional computers at this time.

9.1. Ethical Considerations

In the Age of Industry 4.0, the integration of advanced technologies such as artificial intelligence (AI), machine learning, and the Industrial Internet of Things (IIoT) has become more and more important to strike a balance between technological advancements and ethical considerations [114]. Concerns regarding the security and privacy of personal data are greatly increased by Industry 4.0's growing reliance on data collection and analysis [115]. In this regard, the application of AI algorithms and machine learning techniques increases the likelihood of identifying specific people, potentially endangering their right to privacy [116].

The General Data Protection Regulation (GDPR) is a significant regulatory framework that addresses data collection and analysis and adheres to principles that respect privacy in industry 4.0. It establishes guidelines for the gathering, handling, and archiving of personal data inside the European Union [117]. One thorough framework for the creation and application of AI systems that uphold fundamental rights and are transparent, accountable, and committed to transparency is provided by the European Commission's Ethics Guidelines for Trustworthy AI [118]. Another crucial ethical factor in Industry 4.0 is fairness in AI algorithms and decision-making procedures [119]. AI ethics are seriously threatened by discrimination based on gender, race, or other protected characteristics. Another crucial ethical consideration for Industry 4.0's AI-driven technologies is safety; making it critical to protect people's safety and well-being [120]. Organizations must accept accountability for the choices their AI systems make, as accountability emerges as a fundamental ethical principle, making the design and operation in AI systems vital so that decisions are established [121].

In the framework of industry 4.0, Privacy Enhancing Technologies (PETs) are essential for addressing ethical issues with data protection and privacy as they give people the power to manage and safeguard their personal data when interacting with AI systems [122]. Organizations must create and put into place thorough privacy and data protection policies tailored to the industry 4.0 environment to properly handle these ethical issues [123]. Business prioritized data and algorithm transparency, but there was limited focus on transparency in system development, with varying documentation practices among companies [124].

Employing standards like the IEEE Ethically Aligned Design (EAD) guidelines should be the first step for any organization interested in implementing AI ethics. The stress needed for further work is really important, therefore guidelines are implemented for users to follow [124]. With fairness being important in AI algorithms, organizations must prioritize safety, protect the privacy and security of people's personal information, and hold their AI systems accountable for their decisions [125].

9.2. Limitation

It is important to acknowledge several limitations that could impact the scope and implications of this research. This research has limitations in scope and timeliness, limited to information until September 2023, possibly excluding recent developments in rapidly evolving Industry 4.0 and AI tech. It also might not encompass all AI applications or sector specific intricacies as the focus is on privacy concerns.

10. Conclusion

With the development of AI skyrocketing in many different sectors, Industry 4.0 has caught up on the AI-driven mitigation rules and techniques. The progression of AI in cybersecurity has led to many researches to identify multiple research gaps in different forms. Advanced AI algorithms help in the protection and safety of large and sensitive data, and this is easily implemented using privacy-preserving techniques and can be ruled out using multiple different regulatory frameworks and compliances. The applications of AI present in the vast scheme of cybersecurity were studied through identifying 127 primary field researches out of the 879 linked articles over the past three months. From the observation made, articles in connection with the topics are being published at a great pace, but many factors such as credibility of historical data used, or the association of certain topics should be monitored. This article is directed in a way where the basics of AI and Industry 4.0 are introduced, but as the read continues, advanced links between the topics and compliances are made. Overall, this paper unified a subtle understanding of the challenges faced in Industry 4.0, and with the help of AI services, sensitive information can be safeguarded.

Compliance with ethical standards

Acknowledgments

The authors would like to express sincere appreciation and gratitude to Vellore Institute of Technology for constant support and encouragement and to those who have contributed to the completion of this research.

Disclosure of conflict of interest

The authors declare that they have no competing interests.

Availability of Supporting data.

This research does not include specific supporting data. The information presented is based on a comprehensive review and analysis of relevant literature as cited in the references section.

Funding

The authors would like to acknowledge that this study was conducted without external funding.

Authors' contributions

- T.J. wrote the main manuscript text "Introduction", "Background [2.1] Privacy concerns in Industry 4.0", "Search Strategies" and "First Pillar of Defense". Prepared Figure 1.1 and Table 1.
- A.R.P wrote the main manuscript text "Background [2.2] AI-Driven Data Analytics in Industry 4.0", "Second Pillar of Defense" and "Limitation". Prepared Figure 1.2.
- G.S.R wrote the main manuscript text "Field Research" and "Conclusion".
- A.K. wrote the main manuscript text "Cross-disciplinary Approaches" and "Ethical Considerations".
- S.K. wrote the main manuscript text "Regulatory Frameworks and Compliance" and "Future Trends and Scope"
- A.D jointly supervised the work.

All authors contributed to the main manuscript text "Abstract". All authors reviewed the manuscript.

References

[1] IBM. (2023). What is Industry 4.0 and how does it work https://www.ibm.com/topics/industry-4-0#:~:text=Industry%204.0%20is%20revolutionizing%20the

[2] Habib, M. K., & Chimsom, C. (2019). Industry 4.0: Sustainability and Design Principles. *2019 20th International Conference on Research and Education in Mechatronics (REM)*. https://doi.org/10.1109/REM.2019.8744120

[3] (2019). Smart Factory Cybersecurity Risks and the Future of Manufacturing. *Swivel Secure*, 2019. swivelsecure.com/solutions/manufacturing/manufacturing-is-at-risk-from-cybercrime/

[4] Chaudhuri, A. (2016). Internet of things data protection and privacy in the era of the General Data Protection Regulation. *Journal of Data Protection and Privacy*, 2016, *1*, 64-75.https://www.researchgate.net/publication/313920099_Internet_of_things_data_protection_and_privacy_in_the_era_of_the_General_Data_Protection_Regulation

[5] Arya, V. (2023). Ensuring Data Privacy and Integrity in the Age of Industry 4.0. *Insights2Techinfo*. https://insights2techinfo.com/ensuring-data-privacy-and-integrity-in-the-age-of-industry-4-0/

[6] Alazab, M., Gadekallu, T. R., Su. C. (2022). Guest Editorial: Security and Privacy Issues in Industry 4.0 Applications. *IEEE*, 6326-6329. https://doi.org/10.1109/TII.2022.3164741

[7] Duan, L., & Da Xu, L. (2021). Data Analytics in Industry 4.0: A Survey. Information Systems Frontiers. https://doi.org/10.1007/s10796-021-10190-0

[8] Chae, B., & Olson, D. (2022). Technologies and applications of Industry 4.0: insights from network analytics. *International Journal of Production Research, 60*(12), 3682-3704, https://doi.org/10.1080/00207543.2021.1931524

[9] Cioffi, R., Travaglioni, M., Piscitelli, G., Petrillo, A., & De Felice, F. (2020). Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions. *Sustainability, 12*(2), 492. https://doi.org/10.3390/su12020492.

[10] Nelson, G. S. (2015). Practical Implications of Sharing Data: A Primer on Data Privacy, Anonymization, and De - Identification.https://www.researchgate.net/publication/318866074_Practical_Implications_of_Sharing_Data_A_Primer_on_Data_Privacy_Anonymization_and_De-Identification

[11] Tomashchuk, Oleksandr & Landuyt, Dimitri & Pletea, Daniel & Wuyts, Kim & Joosen, Wouter. (2019). A Data Utility-Driven Benchmark for De- Identification Methods. https://doi.org/10.1007/978-3-030-27813-7_5

[12] Ren, W., Tong, X., Du, J. et al. Privacy Enhancing Techniques in the Internet of Things Using Data Anonymization. *Inf Syst Front* (2021). https://doi.org/10.1007/s10796-021-10116-w

[13] World Economic Forum. (n.d.). Fourth Industrial Revolution. https://www.weforum.org/focus/fourth-industrial-revolution/

[14] Wong, Kok-Seng & Kim, Myung. (2017). Privacy Protection for Data-Driven Smart Manufacturing Systems. *International Journal of Web Services Research, 14*, 17-32. https://doi.org/10.4018/IJWSR.2017070102

[15] National Institute of Standards and Technology. (n.d.). Cybersecurity Framework. https://www.nist.gov/cyberframework/

[16] National Institute of Standards & Technology. (n.d.). Cybersecurity for Smart Manufacturing Systems. https://www.nist.gov/programs-projects/cybersecurity-smart-manufacturing-systems

[17] Runde, D. F., Yayboke, E., Bandura, R., Rubin, N., Hammond, M., & Carter, W. A. (2019, May 21). Beyond Technology: The Fourth Industrial Revolution in the Developing World. *Center for Strategic and International Studies.* https://www.csis.org/analysis/beyond-technology-fourth-industrial-revolution-developing-world

[18] Cybersecurity and Infrastructure Security Agency (CISA). (n.d.). Industrial Control Systems.https://www.cisa.gov/topics/industrial-control-systems

[19] Balbix. (n.d.). Cyber Attack Vectors & How to Avoid Them. https://www.balbix.com/insights/attack-vectors-and-breach-methods/

[20] AFP. (2016). ThyssenKrupp hit by hackers eyeing industrial secrets. https://www.timesofisrael.com/thyssenkrupp-hit-by-hackers-eyeing-industrial-secrets/

[21] Shaban, H., & Nakashima, E. (2017). Pharmaceutical giant rocked by ransomware attack. Washington Post. https://www.washingtonpost.com/news/the-switch/wp/2017/06/27/pharmaceutical-giant-rocked-by-ransomware-attack/

[22] Rivlin, P. (2019, September 26). Iran attacks Saudi Arabia. Iqtisadi: Middle East Economy. https://dayan.org/content/iran-attacks-saudi-arabia

[23] Liu, J., Yuan, C., Lai, Y., & Qin, H. (2020). Protection of Sensitive Data in Industrial Internet Based on Three-Layer Local/Fog/Cloud Storage. *Security and Communication Networks*, 2020, 1–16. https://doi.org/10.1155/2020/2017930

[24] Coos, A. (2022). 5 Ways Big Companies Protect their Data. Endpoint Protector Blog. https://www.endpointprotector.com/blog/5-ways-big-companies-protect-their-data/

[25] Zhu, T., Ye, D., Wang, W., Zhou, W., & Yu, P. (2020). More Than Privacy: Applying Differential Privacy in Key Areas of Artificial Intelligence. *IEEE Transactions on Knowledge and Data Engineering, 34*(6), 2824-2843. https://doi.org/10.1109/TKDE.2020.3014246

[26] Zhu, T., Yu, P. S. (2019). Applying Differential Privacy Mechanism in Artificial Intelligence. *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, 1601-1609. https://doi.org/10.1109/ICDCS.2019.00159

[27] Arachchige, P. Ch. M., Bertok, P., Khalil, I., Liu, D., Camtepe, S., Atiquzzaman, M. (2020). Local Differential Privacy for Deep Learning. *IEEE Internet of Things Journal*, 2020, *7*(7), 5827-5842. https://doi.org/10.1109/JIOT.2019.2952146

[28] Zhao, Y., Zhao, J., Yang, M., Wang, T., Wang, N., Lyu, L., Niyato, D., & Lam, K.-Y. (2020). Local Differential Privacy based Federated Learning for Internet of Things. *IEEE Internet of Things Journal*, 2021, *8*(11), 8836-8853. https://doi.org/10.1109/JIOT.2020.3037194

[29] Hao, M., Li, H., Luo, G., Yang, H., Liu, S. (2019). Efficient and Privacy-Enhanced Federated Learning for Industrial Artificial Intelligence. *IEEE Transactions on Industrial Informatics*, 2020, *16*(10), 6532-6542. https://doi.org/10.1109/TII.2019.2945367

[30] Zhang, C., Xie, Y., Bai, H., Yu, B., Li, W., & Gao, Y. (2021). A survey on federated learning. *Knowledge-Based Systems, 216*, 106775. https://doi.org/10.1016/j.knosys.2021.106775

[31] David, A., Jafar, S., & Alexander, H. (2022). Homomorphic Encryption for Machine Learning and Artificial Intelligence Applications. https://doi.org/10.2172/1886256

[32] Behera, S., Prathuri, J. R. (2020). Application of Homomorphic Encryption in Machine Learning. *2020 2nd PhD Colloquium on Ethically Driven Innovation and Technology for Society (PhD EDITS)*, 2020, 1-2. https://doi.org/10.1109/PhDEDITS51180.2020.9315305

[33] Catak, F. O., Aydin, I., Elezaj, O., & Yildirim-Yayilgan, S. (2020). Practical Implementation of Privacy Preserving Clustering Methods Using a Partially Homomorphic Encryption Algorithm. *Electronics, 9*(2), 229. https://doi.org/10.3390/electronics9020229

[34] Pulido-Gaytan, L. B., Tchernykh, A., Cortés-Mendoza, J. M., Babenko, M., & Radchenko, G. (2021). A Survey on Privacy-Preserving Machine Learning with Fully Homomorphic Encryption. *Communications in Computer and Information Science*,115–129. https://doi.org/10.1007/978-3-030-68035-0_9

[35] Lytvyn, O., & Nguyen, G. (2023). Secure Multi-Party Computation for Magnetic Resonance Imaging Classification. *Procedia Computer Science*, 220, 24–31. https://doi.org/10.1016/j.procs.2023.03.006

[36] Patra, A., Schneider, T., Suresh, A., & Yalame, H. (2021). ABY2.0: New Efficient Primitives for STPC with Applications to Privacy in Machine Learning. https://openreview.net/forum?id=hrV8Tn5w5D0

[37] Augusto, C., Morán, J., Riva, C. D. L., Tuya, J. (2019). Test-Driven Anonymization for Artificial Intelligence. *2019 IEEE International Conference On Artificial Intelligence Testing (AITest)*, 2019, 103-110. https://doi.org/10.1109/AITest.2019.00011

[38] 3AI. (2023). Impact of AI on Data Anonymization. 3AI. https://3ai.in/impact-of-ai-on-data-anonymization/

[39] Hariharan, S. (2023). Leveraging AI-Driven Tokenization and Threat Detection for Data Security. *Protecto*. https://www.protecto.ai/blog/leveraging-ai-driven-tokenization-for-data-security

[40] Enhancing Data Security and Privacy with Tokenization. (n.d.)https://www.segmed.ai/blog/enhancing-data-security-and-privacy-with-tokenization

[41] Desik, A. Role of Data Masking & FHE in Safeguarding Customer Privacy. (n.d.). www.tcs.com. https://www.tcs.com/insights/blogs/data-privacy-masking-personal-information

[42] Zhang, M., Huang, S., Shen, G., & Wang, Y. (2023). PPNNP: A privacy-preserving neural network prediction with separated data providers using multi-client inner-product encryption. *Computer Standards & Interfaces, 84*, 103678. https://doi.org/10.1016/j.csi.2022.103678

[43] El Saj, R., Sedgh Gooya, E., Alfalou, A., & Khalil, M. (2021). Privacy-Preserving Deep Neural Network Methods: Computational and Perceptual Methods—An Overview*. Electronics, 10*(11), 1367. https://doi.org/10.3390/electronics10111367

[44] Ushakov, Y. A., Polezhaev, P. N., Shukhman, A. E., Ushakova, M. v., Nadezhda, M.V. (2018). Split Neural Networks for Mobile Devices. *2018 26th Telecommunications Forum (TELFOR)*, 2018, 420-425. https://doi.org/10.1109/TELFOR.2018.8612133

[45] Bonawitz, K. A., Ivanov, V., Kreuter, B., Marcedone, A., McMahan, H. B., Patel, S., Ramage, D., Segal, A., & Seth, K. (2016). Practical Secure Aggregation for Federated Learning on User-Held Data. https://research.google/pubs/pub45808/

[46] Bonawitz, K., Ivanov, V., Kreuter, B., Marcedone, A., Mcmahan, H., Patel, S., Ramage, D., Segal, A., & Seth, K. (n.d.). Practical Secure Aggregation for Privacy-Preserving Machine Learning. https://eprint.iacr.org/2017/281.pdf

[47] Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and Harnessing Adversarial Examples.https://arxiv.org/abs/1412.6572

[48] NeurIPS. (2023). The Thirty-seventh Annual Conference on Neural Information Processing Systems. https://nips.cc

[49] International Conference on Machine Learning (ICML). (2024). The Forty-first International Conference on Machine Learning. *Messe Wien Exhibition Congress Center, Vienna, Austria*. https://icml.cc

[50] IEEE/CVF. (2024). The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2024. *Seattle Convention Center*. https://cvpr.thecvf.com

[51] ECCV 2022. (n.d.). European Conference on Computer Vision 2022. https://eccv2022.ecva.net

[52] Privacy Preservation - an overview. ScienceDirect Topics (n.d.). *ScienceDirect.*https://www.sciencedirect.com/topics/computer-science/privacy-preservation

[53] Wirth, F. N., Meurers, T., Johns, M., & Prasser, F. (2021). Privacy-preserving data sharing infrastructures for medical research: systematization and comparison. *BMC Medical Informatics and Decision Making, 21*(1). https://doi.org/10.1186/s12911-021-01602-x

[54] Yadav, S., & Tiwari, N. (2023). Privacy preserving data sharing method for social media platforms. *PLOS ONE, 18*(1). https://doi.org/10.1371/journal.pone.0280182

[55] Frąckiewicz, M. (2023). The Role of Secure Enclaves in Modern Cybersecurity. *TS2 SPACE.*https://ts2.space/en/the-role-of-secure-enclaves-in-modern-cybersecurity/#gsc.tab=0

[56] Lomotey, R. K., Kumi, S., & Deters, R. (2022). Data Trusts as a Service: Providing a platform for multi-party data sharing. *International Journal of Information Management Data Insights, 2*(1), 100075. https://doi.org/10.1016/j.jjimei.2022.100075

[57] Ruhaak, A. (2021). How data trusts can protect privacy. MIT. https://www.technologyreview.com/2021/02/24/1017801/data-trust-cybersecurity-big-tech-privacy/

[58] Smit, D., Sunet Eybers, & Smith, J. (2022). A Data Analytics Organisation's Perspective on Trust and AI Adoption. *Communications in Computer and Information Science*, 47–60. https://doi.org/10.1007/978-3-030-95070-5_4

[59] Ducato, R. (2020). Data protection, scientific research, and the role of information. Computer Law & Security Review, 37. *ScienceDirect*. https://doi.org/10.1016/j.clsr.2020.105412

[60] Sartor, G. (2020). The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. https://doi.org/10.2861/293

[61] Mulgund, P., Mulgund, B., Sharman, R., & Singh, R. (2021). The Implications of the California Consumer Privacy Act (CCPA) on Healthcare Organizations: Lessons Learned From Early Compliance Experiences. *Health Policy and Technology, 10*. https://doi.org/10.1016/j.hlpt.2021.100543

[62] Putman, C., Donnelly Mr, S., & Broos, L. (2020). Assessing the Impact of the Implementation of the California Consumer Privacy Act on the United States through Policy Evaluation. https://essay.utwente.nl/85318/

[63] Garlie, M. (2020). California Consumer Privacy Act of 2018: A Study of Compliance and Associated Risk.https://www.proquest.com/openview/e341d9d4ddbdab7a174d21f85d6247f7/1?pqorigsite=gscholar&cbl=18750&diss=y

[64] MINISTRY OF LAW, JUSTICE AND COMPANY AFFAIRS (Legislative Department).https://eprocure.gov.in/cppp/rulesandprocs/kbadqkdlcswfjdelrquehwuxcfmijmuixngudufgbuubgubfugbububjxcgfvsbdihbgfGhdfgFHytyhRtMjk4NzY=

[65] Basu, S., & Jones, R. (2003). E-commerce and the law: a review of India's information technology act, 2000. *Contemporary South Asia, 12*(1), 7–24. https://doi.org/10.1080/0958493032000123344

[66] Bahn, C. (2021). IEEE Standard for Technical Framework and Requirements of Trusted Execution Environment based Shared Machine Learning*. IEEE Computer Society*. https://standards.ieee.org/ieee/2830/10231/

[67] Seaman, M. (n.d.). 802.1X: Port-Based Network Access Control. Revision of IEEE Std 802.1X-2010 and amendments. https://1.ieee802.org/security/802-1x/

[68] (2017). IEEE 802.11i Wireless LAN Security. *BrainKart*https://www.brainkart.com/article/IEEE-802-11i-Wireless-LAN-Security_8486/

[69] (2021). Understanding IEC 62443. https://www.iec.ch/blog/understanding-iec-62443

[70] (2020). IEC 61701:2020. *TC 82 - Solar photovoltaic energy systems, 3*, 1-30. https://webstore.iec.ch/publication/59588

[71] IEC 80001-1:2021. IEC Webstore. (n.d.). *Webstore.iec.ch, 2*, 1-75. https://webstore.iec.ch/publication/34263

[72] ISO. (2022). ISO/IEC 27001:2022. Information security management systems. ISO. https://www.iso.org/standard/27001

[73] (2019). IEC 62645:2019. *IEC Webstore*. (n.d.). 2, 1-112. https://webstore.iec.ch/publication/32904

[74] (2023). IEC 61850:2021 SER. *IEC Webstore*. LVDC. (n.d.). 1, 8186. https://webstore.iec.ch/publication/6028

[75] Chauhan, D. (2023). Data Protection Law in India. ResearchGate. https://www.researchgate.net/publication/370944754_Data_Protection_Law_in_India

[76] Habrat, D. (2020). Legal challenges of digitalization and automation in the context of Industry 4.0. *Procedia Manufacturing, 51*, 938-942. https://doi.org/10.1016/j.promfg.2020.10.132

[77] Taeihagh, A., Ramesh, M., & Howlett, M. (2021). Assessing the regulatory challenges of emerging disruptive technologies. Regulation & Governance. Advance online publication. https://doi.org/10.1111/rego.12392

[78] Zainol, Z., Hassan, K., Wan Hussain, W. M. H., & Phuoc, J. (2019). Adaptive Regulation for Industry 4.0.https://doi.org/10.2991/icss-19.2019.8

[79] Government of Malaysia. (2018). Industry4WRD: National Policy on Industry 4.0. Malaysia Government Portal. https://www.malaysia.gov.my/portal/content/31224

[80] Institution of Engineers Malaysia. (2019). Malaysia national industry 4.0 Policy-Industry4wrd: Opportunities and challenges [PDF]. *Universiti Malaysia Perlis Institutional Repository*. https://rb.gy/vrc4ed

[81] Fruhlinger, J. (2020). Marriott data breach FAQ: How did it happen and what was the impact? CSO.https://www.csoonline.com/article/567795/marriott-data-breach-faq-how-did-it-happen-and-what-was-the-impact.html

[82] Source Defense. (2022). British Airways: A Case Study in GDPR Compliance Failure. Security Boulevard. https://securityboulevard.com/2022/09/british-airways-a-case-study-in-gdpr-compliance-failure/

[83] Raptis, T., Passarella, A., & Conti, M. (2019). Data Management in Industry 4.0: State of the Art and Open Challenges*. IEEE Access*, 1-1. https://doi.org/10.1109/ACCESS.2019.2929296

[84] Scarfone, K., Souppaya, M., & Sexton, M. (2007). NIST SP 800-111: Guide to Storage Encryption Technologies for End User Devices. *National Institute of Standards and Technology (NIST)*. https://csrc.nist.gov/pubs/sp/800/111/final

[85] World Economic Forum. (2018). Data Policy in the Fourth Industrial Revolution: Insights on personal data [PDF]. https://www3.weforum.org/docs/WEF_Data_Policy_in_the_Fourth_Industrial_Revolution_2020.pdf

[86] Khan, M. (2021). Data Minimization—A Practical Approach. ISACA. https://www.isaca.org/resources/news-and-trends/industry-news/2021/data-minimization-a-practical-approach

[87] Onik, M. M. H., Kim, C.-S., & Yang, J. (2019). Personal Data Privacy Challenges of the Fourth Industrial Revolution. *2019 21st International Conference on Advanced Communication Technology (ICACT),* 2019, 635-638. https://doi.org/10.23919/ICACT.2019.8701932

[88] Osborne, C. (2021). Volkswagen, Audi disclose data breach impacting over 3.3 million customers, interested buyers. *ZDNet*.https://www.zdnet.com/article/volkswagen-audi-disclose-data-breach-impacting-over-3-3-million-customers-interested-buyers/

[89] Khaitan, A. (2023). BMW Potential Data Breach Puts Customers Information At Risk! The Cyber Express. https://thecyberexpress.com/bmw-data-breach-customers-information-risk/#:~:text=The%20saga%20of%20BMW%20data,automobile%20company%20suffered%20substantial%20damage

[90] Noss, S. (2023). Data Privacy Training for Employees. *DataGrail*https://www.datagrail.io/blog/data-privacy/data-privacy-training-for-employees/#:~:text=What%20Are%20Data%20Privacy%20and,importance%20of%20protecting%20sensitive%20information

[91] Proietti Franceschilli, C. (2019). Beyond the GDPR: Data protection in the context of industry 4.0 (Master's thesis, Luiss Guido Carli). http://tesi.luiss.it/25791/

[92] Bleich, C. (n.d.). Industry 4.0 And The Future Of Training Workers. https://www.edgepointlearning.com/blog/industry-4-0/

[93] Ometov, A., Bezzateev, S., Mäkitalo, N., Andreev, S., Mikkonen, T., & Koucheryavy, Y. (2018). Multi-Factor Authentication: A Survey. *Cryptography, 2*(1), 1. https://doi.org/10.3390/cryptography2010001

[94] Noss, S. (2023). Data Privacy Training for Employees. *DataGrail*https://www.datagrail.io/blog/data-privacy/data-privacy-training-for-employees/#:~:text=What%20Are%20Data%20Privacy%20and,importance%20of%20protecting%20sensitive%20information

[95] Avdibasic, E., Amanzholova, S., & Durakovic, B. (2022). Cybersecurity challenges in Industry 4.0: A state-of-the-art review. *Defense and Security Studies, 3*, 32-49. https://doi.org/10.37868/dss.v3.id188

[96] Guest (2021). Multi-Factor Authentication and NIST Password Guidelines. *ID R&D*.https://www.idrnd.ai/understanding-multi-factor-authentication-as-related-to-nist/#:~:text=MFA%20adds%20a%20security%20layer

[97] Bernstein, E. (2014). The transparency trap. *Harvard Business Review*. https://hbr.org/2014/10/the-transparency-trap

[98] Government Accountability Office. (2021). SolarWinds Cyberattack Demands Significant Federal and Private-Sector Response. https://www.gao.gov/blog/solarwinds-cyberattack-demands-significant-federal-and-private-sector-response-infographic

[99] European Commission. (2016). Data protection in the EUhttps://commission.europa.eu/law/law-topic/data-protection/data-protection-eu_en

[100] Waldo, J., Lin, H. S., Millett, L. I. (Eds.). (2007). Engaging Privacy and Information Technology in a Digital Age. *Committee on Privacy in the Information Age, National Research Council*. http://www.nap.edu/catalog/11896.html

[101] Maggiolino, M., & Zoboli, L. (2021). The Intersection Between Intellectual Property and Antitrust Law. In I. Calboli & M. L. Montagnani (Eds.), *Handbook on Intellectual Property Research*.https://doi.org/10.1093/oso/9780198826743.003.0009

[102] Prinsloo, J., Sinha, S., & von Solms, B. (2019). A Review of Industry 4.0 Manufacturing Process Security Risks. *Applied Sciences, 9*(23), 5105. https://doi.org/10.3390/app9235105

[103] Safdar, N. M., Banja, J. D., & Meltzer, C. C. (2020). Ethical considerations in artificial intelligence. *European Journal of Radiology, 122*, 108768. https://doi.org/10.1016/j.ejrad.2019.108768

[104] Mou, Y., & Xu, K. (2017). The media inequality: Comparing the initial human-human and human-AI social interactions. *Computers in Human Behavior, 72*, 432-440. https://doi.org/10.1016/j.chb.2017.02.067

[105] Lai, Y., Kankanhalli, A., & Ong, D. (2021). Human-AI Collaboration in Healthcare: A Review and Research Agenda. https://doi.org/10.24251/HICSS.2021.046

[106] Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. *International Journal of Information Management, 53*, 102104. https://doi.org/10.1016/j.ijinfomgt.2020.102104

[107] Schranz, M., Umlauft, M., Sende, M., & Elmenreich, W. (2020). Swarm Robotic Behaviors and Current Applications. *Frontiers in Robotics and AI*, 7. https://doi.org/10.3389/frobt.2020.00036

[108] Shahzad, M. M., Saeed, Z., Akhtar, A., Munawar, H., Yousaf, M. H., Baloach, N. K., & Hussain, F. (2023). A Review of Swarm Robotics in a NutShell. *Drones, 7*(4), 269. https://doi.org/10.3390/drones7040269

[109] Clark, S. (2021). 4 reasons why explainable AI is the future of AI. Digital Experience.https://www.cmswire.com/digital-experience/4-reasons-why-explainable-ai-is-the-future-of-ai/

[110] Dégallier-Rochat, S., Kurpicz-Briki, M., Endrissat, N., & Yatsenko, O. (2022). Human augmentation, not replacement: A research agenda for AI and robotics in the industry. *Frontiers in Robotics and AI*, 9. https://doi.org/10.3389/frobt.2022.997386

[111] Reitmayr, G., Langlotz, T., Wagner, D., Mulloni, A., Schall, G., Schmalstieg, D., & Pan, Q. (2010). Simultaneous Localization and Mapping for Augmented Reality (PDF). *International Symposium on Ubiquitous Virtual Reality*, 5-8. https://doi.org/10.1109/ISUVR.2010.12

[112] Kanwar, P. (2023). Quantum Computing in Artificial Intelligence Around the Corner. AI & Machine Learning. https://www.einfochips.com/blog/quantum-computing-in-artificial-intelligence-around-the-corner/

[113] Sarker, I. H. (2022). AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems. *SN COMPUT. SCI., 3*, 158. https://doi.org/10.1007/s42979-022-01043-x

[114] Bajic, B., Rikalovic, A., Suzic, N., & Piuri, V. (2021). Industry 4.0 Implementation Challenges and Opportunities: A Managerial Perspective. *IEEE Systems Journal, 15*, 546-559. https://doi.org/10.1109/JSYST.2020.3023041

[115] Manheim, K., & Kaplan, L. (2019). Artificial Intelligence: Risks to Privacy and Democracy. *21 Yale J.L. & Tech, 106*. https://yjolt.org/sites/default/files/21_yale_j.l._tech._106_0.pdf

[116] Rhahla, M., Allegue, S., & Abdellatif, T. (2021). Guidelines for GDPR compliance in Big Data systems. *Journal of Information Security and Applications, 61*, 102896. https://doi.org/10.1016/j.jisa.2021.102896

[117] Stix, C. (2021). Actionable Principles for Artificial Intelligence Policy: Three Pathways. *Science and Engineering Ethics, 27*(1), 15. https://doi.org/10.1007/s11948-020-00277-3

[118] Pessach, D., & Shmueli, E. (2020). Algorithmic fairness. *arXiv*. https://doi.org/10.48550/arXiv.2001.09784

[119] Keshta, I. (2022). AI-driven IoT for smart health care: Security and privacy issues. *Informatics in Medicine Unlocked*, *30*, 100903. https://doi.org/10.1016/j.imu.2022.100903

[120] Doshi-Velez, F., Kortz, M., Budish, R., Bavitz, C., Gershman, S., O'Brien, D., Scott, K., Schieber, S., Waldo, J., Weinberger, D., Weller, A., & Wood, A. (2019). Accountability of AI Under the Law: The Role of Explanation. https://doi.org/10.48550/arXiv.1711.01134

[121] Cha, S.-C., Hsu, T.-Y., Xiang, Y., & Yeh, K.-H. (2018). Privacy Enhancing Technologies in the Internet of Things: Perspectives and Challenges. *IEEE Internet of Things Journal*, 2019, *6*(2), 2159-2187.https://doi.org/10.1109/JIOT.2018.2878658

[122] Bygrave, A. L. (2010). Privacy and Data Protection in an International Perspective. *Scandinavian Studies in Law*, 56, 165-200. https://scandinavianlaw.se/pdf/56-8.pdf

[123] Vakkuri, V., Kemell, K. -K., Kultanen, J., & Abrahamsson, P. (2020). The Current State of Industrial Practice in Artificial Intelligence Ethics. *IEEE Software, 37*(4), 50-57. https://doi.org/10.1109/MS.2020.2985621

[124] Novelli, C., Taddeo, M., & Floridi, L. (2023). Accountability in artificial intelligence: What it is and how it works. AI & Soc. https://doi.org/10.1007/s00146-023-01635-y