# REMINDER

**REMINDER:**

pRivacy-prEserving Machine LearnIng through secure managemeNt of Data's lifecyclE in distRibuted systems

**Deliverable number: D1.1**

*Use case definition and initial requirements analysis*

chist-era

FWF Austrian Science Fund

UKRI Engineering and Physical Sciences Research Council

uefiscdi Unitatea Executivă pentru Finanțarea Învățământului Superior, a Cercetării, Dezvoltării și Inovării

AGENCIA ESTATAL DE INVESTIGACIÓN

| | |
|---|---|
| **Project Acronym:** | REMINDER |
| **Project Full Title:** | pRivacy-prEserving Machine LearnIng through secure manageMeNt of Data's lifecyclE in distRibuted systems |
| **Call:** | Security and Privacy in Decentralised and Distributed Systems (SPiDDS). 2022 |
| **Grant Number:** | PCI2023-145989-2 |
| **Project URL:** | https://ants.inf.um.es/en/reminder |
| **Editor:** | UMU |
| **Deliverable nature:** | Report |
| **Dissemination level:** | Public |
| **Delivery Date:** | 31/07/2024 |
| **Authors:** | UMU, SIE, UWE, AIT |

*Table 1: Project details.*

# Abstract

The increasing integration of Artificial Intelligence (AI) is driving transformative advancements across various sectors by enabling real-time data-driven decision-making. However, these innovations bring significant challenges related to data security and privacy. The REMINDER project proposes a novel framework to address these concerns in distributed systems by leveraging Federated Learning (FL). This decentralized approach enables collaborative Machine Learning (ML) model training while preserving data privacy. Designed to accommodate resource-constrained devices, the architecture ensures robust security and privacy throughout the data lifecycle. REMINDER addresses diverse security threats, implementing protocols to authenticate legitimate participants in the collaborative training process. This document outlines the conceptual framework of REMINDER's edge-based architecture, exploring its validation through use cases in healthcare and smart buildings.

# Table of Contents

# List of Figures

# List of Tables

# List of Abbreviations

# 1 Introduction

The rapid digitalization and growing adoption of connected devices have emphasized the need for secure and decentralized machine learning approaches. Federated Learning (FL) has emerged as a promising solution to train AI models while preserving privacy by keeping raw data localized. However, applying FL in real-world scenarios introduces challenges related to data security, integrity, and interoperability, particularly in sensitive domains such as healthcare and smart buildings.

AI plays a crucial role in technological advancements, supporting smart cities and eHealth services while contributing to the United Nations' Sustainable Development Goals (SDGs). IoT and 5G/6G networks facilitate real-time data exchange, enabling data-driven decision-making. However, traditional centralized AI architectures raise privacy concerns as data must be shared with central servers. FL mitigates these risks by enabling collaborative model training on local data, ensuring that only model updates are shared instead of raw information. This enhances privacy and fosters collaboration across organizations while maintaining regulatory compliance, such as with the European Union Artificial Intelligence Act.

Despite its advantages, FL faces security risks, including model poisoning and inference attacks that can compromise data privacy. The decentralized nature of FL also poses challenges in authenticating participants and ensuring model integrity. Addressing these risks requires robust techniques such as Differential Privacy (DP), cryptographic protocols, and resilient aggregation methods.

The REMINDER project aims to address these challenges by designing an FL framework that integrates advanced security mechanisms, including encrypted model updates, digital signatures, and robust aggregation techniques. By focusing on two key use cases—eHealth and smart buildings—this project explores FL's potential for enabling privacy-preserving, collaborative learning. This document outlines the project's key concepts, selected use cases, and the technical and security requirements necessary for the successful implementation of the REMINDER framework considering the mentioned use cases.

# 2 Properties and Challenges of Federated Learning

Federated Learning (FL) has emerged as a transformative approach in Machine Learning (ML), particularly in privacy-sensitive domains such as healthcare and smart infrastructure management. By enabling decentralized training, FL minimizes the risk of direct data exposure, aligning with stringent data protection regulations such as GDPR [3]. The inherent privacy benefits of FL are often enhanced by mechanisms like Differential Privacy (DP), which introduces controlled noise to model updates to obscure individual contributions [4]. This decentralized nature also enhances scalability by distributing computational tasks across multiple nodes, thereby reducing dependence on centralized infrastructure.

However, FL faces significant challenges in real-world deployments. One of the most pressing issues is security, as adversaries can exploit vulnerabilities in the training process to conduct poisoning attacks, inference attacks, or compromise model integrity [5]. Furthermore, privacy risks remain a concern, as exposed model updates can be leveraged for data reconstruction or membership inference [6]. In FL, ensuring robust authentication is crucial to prevent adversarial participants from corrupting the global model, particularly in environments with dynamic client participation.

Another key challenge is handling non-iid (non-independent and identically distributed) data. Many FL implementations assume uniform data distribution across clients, but in practice, datasets differ significantly across participants. This issue leads to client drift, where local models diverge from the global objective, reducing performance and slowing convergence [7]. Addressing these heterogeneities requires advanced aggregation functions capable of mitigating the impact of biased updates while preserving accuracy.

These challenges are particularly relevant in the eHealth and Smart Buildings use cases of the REMINDER project. In eHealth, protecting sensitive patient data while enabling collaborative learning is paramount, as security breaches in medical environments could expose confidential health records and disrupt healthcare operations. Similarly, in Smart Buildings, ensuring that energy optimization models do not leak private occupancy data or behavioral patterns is critical for both compliance and user trust.

The REMINDER project directly addresses these concerns by integrating cryptographic techniques, robust authentication mechanisms, and privacy-preserving aggregation functions. In the following sections, we explore how REMINDER mitigates these security and privacy challenges in each use case, ensuring reliable and privacy-conscious federated learning deployments. In the deliverable 4.1 there is a more complete description of the threats involving FL. Specifically, in Table 2 we show the attacks which will be mitigated by our framework.

*Table 2: Attacks mitigated by REMINDER and their corresponding mitigation techniques.*

| Type of Attack | Attack Category | Mitigation Technique in REMINDER |
|---|---|---|
| **Poisoning Attacks** | Security | Robust aggregation (e.g., FFT) |
| **Model Poisoning** | Security | Robust aggregation (e.g., FFT) |
| **Sybil Attacks** | Security | Authentication protocol with Elliptic Curve Cryptography (ECC) |
| **Inference Attacks** | Privacy | Differential Privacy (DP) |
| **Eavesdropping** | Privacy | Cryptography, eg: Fully Homomorphic Encryption (FHE) |

# 3 REMINDER Architecture

The REMINDER project presents a privacy-preserving and decentralized Federated Learning (FL) framework designed to meet the requirements established in this document, it will be explained in detail in Deliverable 2. This architecture ensures compliance with key principles, including:

**Communication and Data Transmission** – Secure, efficient, and authenticated data exchange across all system components.

**Privacy and Security** – Protection of sensitive information through data obfuscation, confidentiality-preserving mechanisms, and secure model aggregation.

**Scalability and Interoperability** – Seamless integration with diverse smart building infrastructures, including modern and legacy systems, while supporting large-scale deployments.

**Federated Learning and Energy Optimization** – Adaptive, robust FL models capable of optimizing energy consumption dynamically while maintaining system efficiency.

**Regulatory Compliance and Auditability** – Alignment with privacy regulations and the implementation of transparent monitoring and auditing mechanisms.

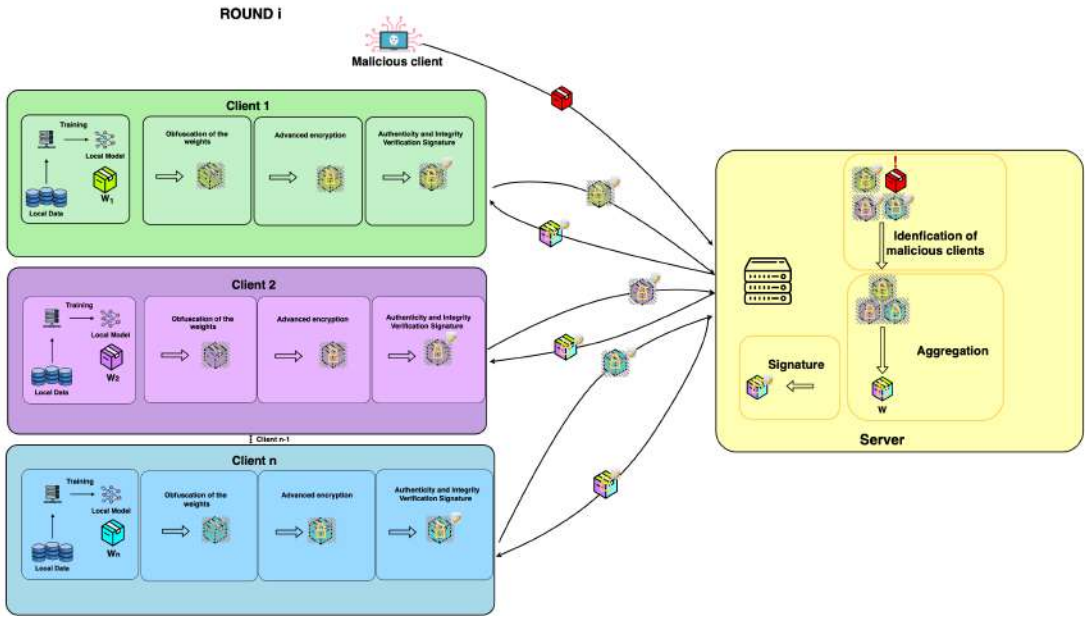


*Figure 1: REMINDER Architecture*

Our proposed architecture is presented in Fig. 1. In this case, during the communication between clients and server we add privacy-preserving techniques such us DP, encrypt methods and signature processes. And then, in the server, it is applied a different aggregation functions and frameworks for detecting malicious clients. In particular, the procedure of REMINDER will be:

**(At clients' side)**

1. We train the model in the "Trainer", a device which is technically capable of this task
2. After the training, the clients will apply DP techniques to obfuscate the weights.

3. The clients will cipher to add a higher degree of privacy. These updates are encrypted using a cipher. This ensures that even if intercepted by a malicious entity, it would be significantly more challenging to infer any sensitive information.
4. The model updates are signed to ensure their integrity and authenticity. The model updates are then sent to the server,

   **(At servers' side)**

5. The server performs a malicious client detector to make a first filter in order to remove as many malicious clients as possible.
6. The server aggregates all updates received from the clients using a robust aggregation function that discards malicious updates. The aggregation should be performed on encrypted data, ensuring that the server never has access to the actual weight values.

This process results in a global model that combines the knowledge from all legitimate clients, ensuring good performance while preserving the privacy of data from all sources. The server signs the updated global model to guarantee its provenance and integrity. The update is then sent to the known clients in its encrypted form, as the weights have not been decrypted at any stage of the process.

# 4 Use case 1 - eHealth

Recent technological breakthroughs have revolutionized medical practices, driving a transformative shift in the healthcare industry. The surge in computational capabilities and data storage has fueled a wave of innovations that now underpin modern healthcare solutions. These advancements have positively impacted patient outcomes and, together with other factors, have supported the steady increase in global life expectancy [8].

Although biomedical data is abundant, its circulation is tightly restricted due to ethical concerns surrounding patient privacy. Regulations such as the General Data Protection Regulation (GDPR) in Europe and the Health Insurance Portability and Accountability Act (HIPAA) in the US impose strict limitations on data access to safeguard patient confidentiality. While these measures are fundamentally justified, they pose significant challenges to AI-driven innovations that have the potential to enhance clinical practices and improve healthcare systems. Using patient data for research or product development requires informed consent, but this consent typically applies only to the initial purpose, preventing further use in subsequent projects. As a result, meeting the data requirements for training robust deep learning models becomes increasingly difficult, as assembling large-scale datasets now often relies on prospective data collection aligned with privacy regulations.

Konecny et al. [9] introduced a machine learning approach that enables model training across decentralized devices without requiring data to leave hospital firewalls. In this setup, multiple data sites operate as nodes, each independently training a deep learning model using their local databases, eliminating the need for data externalization. These decentralized devices communicate solely with a central server, which collects model gradients — rather than raw data — and returns updated parameters derived from the aggregated error signals. Unlike traditional machine learning methods that rely on pooling data from various institutions to enhance model generalization, federated learning offers a promising alternative by developing robust models in a decentralized manner. This approach effectively addresses regulatory challenges tied to sharing sensitive healthcare information.

Despite the well-documented advantages of Federated Learning (FL), it is also acknowledged that machine learning models, including those trained in decentralized settings, remain vulnerable to adversarial interference. Malicious actors can intentionally inject false data during training to deceive the model — a tactic known as data or model poisoning. In FL, attackers may compromise the system by submitting corrupted model updates, ultimately degrading the overall model's performance. Beyond poisoning attacks, recent research highlights additional threats to data privacy in FL environments, such as Membership Inference Attacks (MEIA) [10], Property Inference Attacks (PIA) [11] and Data Reconstruction Attacks (DRA) [12].

The REMINDER framework strengthens its federated learning (FL) architecture with multiple security layers. To prevent adversarial manipulation, digital signatures verify the authenticity of model updates. It also employs advanced aggregation methods, such as Fast Fourier Transform (FFT)-based techniques, to identify and filter out suspicious updates, mitigating the risk of model poisoning. To safeguard data privacy, REMINDER integrates differential privacy by adding controlled noise to model gradients, ensuring that sensitive data patterns could not be inferred while preserving the overall learning process.

## 4.1 eHleath use-cases

To showcase REMINDER's potential in tackling the challenges of developing AI-driven healthcare solutions, we will explore two specific use cases over the course of the project.

### Atrial Fibrillation

Data analytics, cloud computing, and machine learning (ML), have become vital tools for early disease detection. A key application is heart monitoring, where these devices help identify atrial fibrillation — a prevalent cardiac arrhythmia linked to irregular and rapid heartbeats, often leading to strokes. To detect such conditions, wearable sensors or medical equipments capture data like electrocardiograms (ECGs), which are then analyzed using ML algorithms. However, the collection, sharing, and processing of sensitive health data raise critical privacy and security concerns, as any breach could compromise patient confidentiality.

### Aortic Coarctation

Aortic coarctation is a congenital heart defect characterized by a narrowing of the aortic media, reducing the aortic lumen. It accounts for 5 to 8% of all congenital heart disease cases. Diagnosis typically relies on assessing the pressure gradient across the coarctation, which can be estimated using cuff-derived pressure differences between upper and lower extremities, Doppler-based measurements via the Bernoulli equation, or catheter-based measurements — the latter being the gold standard.

To offer a non-invasive alternative to catheterization, [13] introduced personalized blood flow simulations using computational fluid dynamics (CFD) to estimate the pressure drop. However, this method has notable drawbacks: (1) anatomical model reconstruction from medical images requires semi-automated steps with manual adjustments, and (2) CFD simulations are computationally expensive and time-consuming.

A promising solution is to use a reduced set of fully automated anatomical measurements (see figure 2) and leverage deep learning models to predict CFD outputs. Nevertheless, training these models requires access to patient-derived anatomical data, which is subject to strict regulatory controls designed to safeguard patient privacy.

## 4.2 Requirements

The development of AI models for healthcare is often hindered by strict regulations surrounding data privacy. Federated Learning (FL) offers a solution by enabling collaborative model training across multiple hospitals without the need to share sensitive patient data. This section outlines the requirements for implementing a secure, privacy-preserving federated learning system, incorporating additional privacy techniques such as differential privacy and secure aggregation.

### 4.2.1 Functional Requirements

**Federated Learning Setup**

- The system must support distributed model training across multiple hospital nodes.

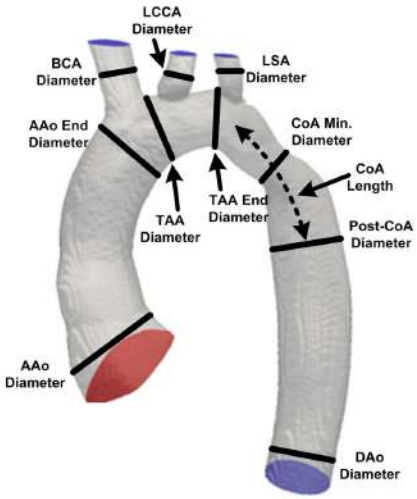


*Figure 2: Reduced set of anatomical features that can be leveraged by a DL algorithm.*

- Each node must train a local model using its own data without exporting any raw data.
- A central server must aggregate local model updates to produce a global model.
- The aggregation method should support partial participation, allowing nodes to contribute asynchronously.

### 4.2.2 Non-Functional Requirements

**Performance**

- While privacy preservation techniques are crucial for protecting the sensitive data, they must not come at the cost of model performance.
- Aggregation latency should be small to support near real-time training.

**Scalability**

- The architecture must support dynamic addition or removal of nodes without retraining the entire model.
- The aggregation process must be optimized to handle increasing data sizes and model complexities.

**Reliability**

- The system must implement fault-tolerant mechanisms, ensuring model training continues even if some nodes drop out.
- Checkpointing should be enabled, allowing model recovery in case of failures.

**Security**

- Nodes must authenticate with the central server using digital signatures to prevent unauthorized participation.
- Encrypted logs must be maintained to track model updates, aggregations, and access requests.

**Privacy**

- Privacy-preserving audit trails must be maintained without revealing sensitive patient information.
- The system must incorporate differential privacy by adding controlled noise to model gradients before they are shared, ensuring that individual patient data cannot be inferred.
- Privacy budgets must be configurable to balance model accuracy and privacy guarantees.

**Regulatory and Compliance Requirements**

- The system must comply with GDPR, HIPAA, and other relevant regulations by ensuring no raw patient data is shared or stored outside hospital premises.
- Patient data must remain anonymized at all stages of training and aggregation.

Implementing federated learning combined with advanced privacy-preserving techniques offers a robust solution for training AI models on decentralized healthcare data. This approach not only ensures compliance with strict data protection laws but also fosters collaboration among hospitals, ultimately leading to more accurate and generalizable AI models. The outlined requirements will guide the development of a secure, scalable, and efficient federated learning system tailored to the unique demands of the healthcare sector.

# 5 Use case 2 - Smart Buildings

The integration of Internet of Things (IoT) devices into smart buildings has transformed the way energy consumption is managed, enabling more efficient use of resources and improved environmental control. These intelligent systems continuously monitor parameters such as temperature, humidity, and occupancy, dynamically adjusting heating, ventilation, and air conditioning (HVAC) operations to optimize comfort while minimizing energy waste. However, the pervasive collection of such data introduces significant privacy and security concerns, particularly when energy usage patterns are correlated with occupant behaviors, potentially revealing sensitive personal information such as daily routines, presence or absence in specific rooms, and even activities performed within a space [14].

With the growing demand for intelligent energy management solutions, Federated Learning (FL) has emerged as a promising approach for enabling privacy-preserving machine learning in smart buildings. Unlike traditional centralized models that require raw data to be transmitted to a central server, FL allows individual buildings to train local models on their own energy consumption data, sharing only aggregated model updates rather than exposing raw measurements. This decentralized paradigm significantly mitigates privacy risks while still enabling collaborative learning, thereby improving predictive accuracy and facilitating more effective energy management strategies [15].

A major challenge in deploying FL across multiple buildings is the heterogeneity in available sensor data. Some buildings are fully equipped with high-resolution IoT infrastructures capable of tracking detailed energy consumption patterns, HVAC operations, weather fluctuations, and motion detection, while others have minimal sensing capabilities, making it difficult to achieve uniform model training across diverse environments. To overcome this limitation, the use of Federated Transfer Learning (FTL) could serve as a promising solution by combining FL with transfer learning techniques to facilitate knowledge transfer from well-instrumented buildings to those with limited data availability [15].

The Pleiades building at the University of Murcia will be used to obtain extensive real-world energy consumption data, including HVAC operations, indoor temperature variations, external weather conditions, and occupant movement [1]. By leveraging this setting, federated models can be pre-trained on buildings with extensive instrumentation and later adapted to new buildings through domain adaptation techniques. Furthermore, through the use of FTL, clustering techniques could be employed to group buildings with similar energy consumption patterns, allowing a representative model to be trained on buildings with sufficient data before being transferred to others with limited IoT coverage [16]. The methodology ensures that knowledge from multiple sources is fused, including real-world datasets, external weather APIs, and smart meter data, thereby enhancing the robustness and generalizability of FL models.

## 5.1 Use of Pleiades BMS

The Pleiades Building Management System (BMS) manages a variety of devices, including air conditioning units, temperature and humidity sensors, and lighting systems, transmitting data in JSON format. This integration allows the building to optimize energy consumption dynamically, responding to real-time environmental and occupancy changes. The monitoring and control infrastructure consists of three-phase and single-phase smart meters such as Qubino 3-Phase Smart Meter, Aeotec Home Energy Meter - Gen5, and WiDom Energy Driven Switch C version, providing granular energy consumption insights. Additionally, Z-Wave sensors like MCOHome


*Figure 3: Illustration of the buildings included in the PLEIAData dataset [1].*

MH9 measure indoor temperature and humidity, while HVAC units, specifically Toshiba VRF, are regulated using Intesis Box WMP – Universal IR smart controllers. A Raspberry Pi 3 Model B+ with an Aeotec Z-Stick Gen5 USB module serves as the gateway for monitoring and control, facilitating interoperability between various IoT devices.

A notable challenge in integrating the Pleiades BMS involved actuation support and minimizing latency in control operations. Indeed, delays in reflecting actuation changes, particularly in HVAC systems, could be identified as a bottleneck. These delays could arise due to the hierarchical nature of HVAC control, where multiple splits within a building are managed by a single gateway, leading to cyclic access times for individual unit updates. To address this, an asynchronous communication strategy will be implemented, significantly improving response times and ensuring real-time synchronization of environmental control parameters. Furthermore, external environmental data from the Instituto Murciano de Investigación y Desarrollo Agrario y Medioambiental (IMIDA) [1] and the Sistema de Información Agrometeorológica de la Región de Murcia (SIAM)[2] will be integrated into the system. These sources provide continuous weather-related insights, enhancing predictive energy models and adaptive control strategies.

This advanced IoT-enabled infrastructure underscores the potential of FL for privacy-preserving energy optimization, as data from hundreds of IoT devices can be leveraged without exposing sensitive user patterns. The collected data will form a crucial benchmark for training FL models, supporting adaptive and scalable energy management frameworks in diverse smart building environments.

## 5.2 Risks and mitigation for the use case

Despite the benefits of FL in preserving data privacy, there remain inherent risks in the potential misuse of energy analytics to infer sensitive information about building occupants. When en-

[1] https://www.imida.es/
[2] http://siam.imida.es/apex/f?p=101:1:2690254026345340::NO:::

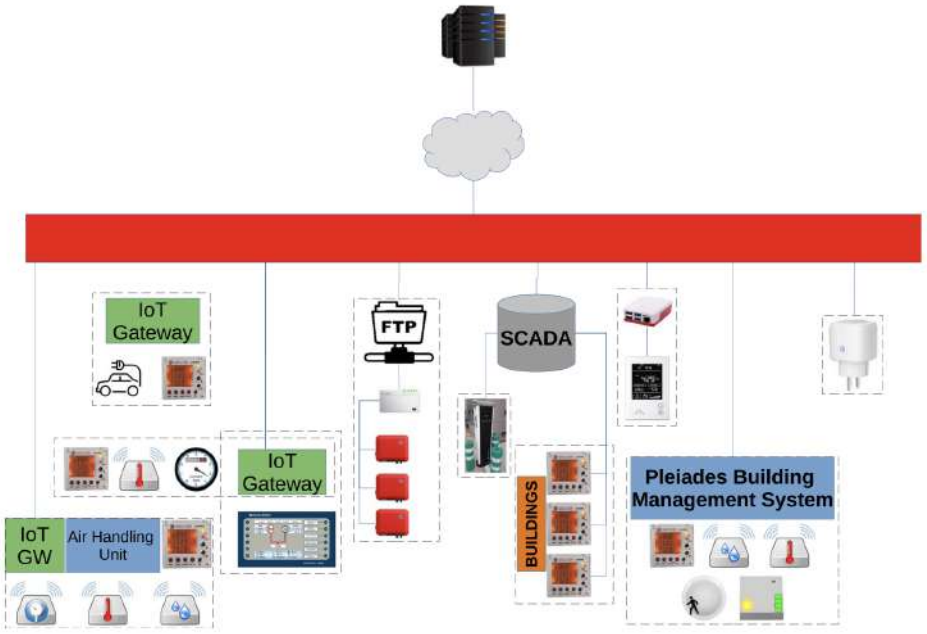


*Figure 4: Architecture diagram UMU Pilot26.[2]*

ergy consumption data is cross-referenced with external sources such as social media check-ins, CCTV footage, or WiFi login records, it becomes possible to de-anonymize individuals and reveal behavioral patterns. For example, if an individual posts on social media that they are "leaving work" and an immediate drop in energy usage is detected in their office, an attacker could correlate this information to infer specific movements and routines [14]. Beyond de-anonymization, energy consumption anomalies can also be exploited to reveal confidential activities. Increased nighttime energy usage might indicate that an employee is working on a highly sensitive project, while irregular HVAC patterns in residential spaces could expose personal habits, such as the use of medical devices or shower schedules, which could be misused for targeted advertising or social engineering attacks. The ability to track occupancy trends over time further exacerbates security risks, as it enables adversaries to predict when buildings are most vulnerable to physical security threats such as burglary or unauthorized access.

To counter these threats, the REMINDER framework incorporates multiple layers of security mechanisms into its FL architecture. Digital signatures ensure that model updates originate from authenticated nodes, preventing adversarial tampering during transmission. Additionally, REMINDER employs robust aggregation strategies such as Fast Fourier Transform (FFT)-based aggregation (FedRDF), which filters out anomalous updates to prevent poisoning attacks and adversarial interference in the global model. Differential privacy techniques are also embedded in the FL training process to prevent reverse-engineering of individual energy usage patterns. Controlled noise is added to model gradients before updates are transmitted, ensuring that no single building's contribution can be distinguished, while still allowing meaningful learning across the federated network.

A significant challenge in the deployment of AI-driven energy management solutions is the integration of legacy BMSs, which were not designed for modern IoT interoperability. Many commercial and residential buildings continue to operate on outdated infrastructure that lacks native support for real-time data exchange, making it difficult to implement advanced federated learning models. To address this, the main goal is to introduce an interoperability layer that

bridges legacy systems with scalable, federated AI-driven energy management platforms [17]. This framework enables protocol translation, allowing legacy communication standards such as Modbus, BACnet, and proprietary BMS protocols to be mapped into IoT-compatible formats like FIWARE NGSI-LD. Smart edge computing nodes and actuators serve as intermediaries where direct integration is not feasible, enabling older buildings to participate in FL without requiring a full infrastructure overhaul. Additionally, automated demand response (ADR) strategies could also be implemented, leveraging FL insights to dynamically adjust HVAC operations, lighting controls, and energy storage usage. This ensures that even buildings with legacy systems can achieve energy optimization without manual intervention, reducing energy costs while maintaining occupant comfort.

## 5.3 Requirements for Smart Buildings use case

The Smart Buildings use case within REMINDER requires a framework that ensures energy efficiency, privacy, security, and interoperability while supporting heterogeneous infrastructures. The following requirements outline the key functional, security, and scalability aspects necessary for effective federated learning (FL) deployment in smart buildings.

**Communication and Data Transmission**

- IoT devices must communicate with the edge server using secure, encrypted channels to prevent unauthorized access to data.
- All IoT devices and edge nodes must authenticate themselves before transmitting data to ensure only legitimate participants contribute to the system.
- The system must support low-latency and efficient data transmission, enabling real-time energy optimization and federated learning updates.
- Communication protocols must be resilient to network disruptions, allowing nodes to synchronize efficiently even in environments with intermittent connectivity.
- Mechanisms must be in place to detect and mitigate data inconsistencies caused by transmission errors or device failures.

**Privacy and Security**

- Sensitive data must be obfuscated or anonymized before being shared or used in FL to protect privacy.
- Model updates must be protected against inference attacks, ensuring that no sensitive building information can be reconstructed.
- The system must ensure that only legitimate and authenticated updates are included in FL aggregation.
- Secure aggregation techniques must be applied to ensure that federated learning processes preserve privacy and data integrity.
- Devices must be capable of securely storing and processing data, ensuring that sensitive information is protected from unauthorized local access.
- The system must provide mechanisms for detecting and isolating compromised nodes, preventing malicious contributions from affecting federated training.

- Authentication and authorization mechanisms must be enforced at all system layers to prevent unauthorized access to FL models, data, and communication channels.
- There must be mechanisms to allow privacy-preserving collaboration between multiple stakeholders, ensuring that different organizations or entities can share insights while maintaining data sovereignty.

**Scalability and Interoperability**

- The FL architecture must support a growing number of buildings without performance degradation.
- Communication overhead must be minimized to optimize bandwidth usage, particularly in environments with limited network connectivity.
- The system must support interoperability with existing Building Management Systems (BMS), allowing legacy systems to integrate without major infrastructure modifications.
- The framework must provide modular and extensible interfaces, allowing new smart building technologies and protocols to be incorporated without requiring fundamental changes to the system.
- The architecture must be compatible with heterogeneous hardware across different buildings, ensuring that FL training can be performed on diverse computational resources.
- Mechanisms must be in place to optimize computational load balancing, distributing federated training tasks efficiently across different buildings to prevent bottlenecks.
- Energy management strategies must be adaptable to different regulatory and operational constraints, allowing custom policies to be applied based on geographic, economic, or user-defined constraints.

**Federated Learning and Energy Optimization**

- FL models must adapt dynamically based on occupancy patterns, external conditions, and energy demand variations.
- The system must support heterogeneous sensor deployments, allowing buildings with different levels of instrumentation to participate in federated training.
- The aggregation function must be resilient to adversarial updates, ensuring the integrity and reliability of the federated model.
- The FL models must be capable of continuous learning, improving performance over time by integrating new data while avoiding catastrophic forgetting.
- Model updates must be optimized for energy efficiency, ensuring that federated training does not introduce excessive computational overhead on resource-constrained devices.
- There must be mechanisms to evaluate model drift and performance degradation, ensuring that FL models remain accurate and effective over time.
- The system must allow for cross-building collaboration, enabling federated learning across different facilities without requiring raw data exchange.
- Energy optimization models must incorporate predictive analytics, allowing buildings to anticipate demand fluctuations and proactively adjust energy usage.

**Regulatory Compliance and Auditability**

- All data processing must comply with GDPR and relevant data protection laws.
- The system must include auditability mechanisms to enable transparency in energy optimization decisions while maintaining data confidentiality.
- There must be a clear framework for accountability, ensuring that stakeholders can verify compliance with data protection and security regulations.
- The system must provide detailed logging and monitoring, allowing for real-time auditing and historical analysis of energy optimization processes.
- FL models must be designed with explainability and interpretability in mind, ensuring that predictions and recommendations can be justified to stakeholders.

By meeting these requirements, the REMINDER framework is intended to ensure that smart buildings can efficiently leverage FL for energy optimization while maintaining privacy, security, and interoperability.

# 6 Conclusion and Future Directions

The REMINDER project is focused on designing a privacy-preserving FL framework tailored for decentralized environments. By defining key security and privacy requirements, this project lays the groundwork for a robust FL approach that can be applied in diverse domains. The identified use cases, namely eHealth and smart buildings, serve as testbeds for evaluating the feasibility of FL in real-world applications, considering challenges such as data heterogeneity, security risks, and interoperability with existing infrastructures.

As future activities, the project will explore methodologies to optimize the efficiency, scalability, and resilience of the FL framework. Future work will focus on refining security mechanisms, implementing privacy-preserving techniques, and conducting experimental validation in simulated and real-world environments. Additionally, the framework will be designed to accommodate emerging challenges in federated AI, enabling its adaptation to other domains such as finance, smart cities, and industrial automation. The REMINDER project aims to set the foundation for secure, scalable, and efficient federated learning solutions that balance innovation with privacy protection.

# References

[1] A. M. Ibarra, A. González-Vidal, and A. Skarmeta, "PLEIAData: consumption, HVAC, temperature, weather and motion sensor data for smart buildings applications," *Scientific Data*, vol. 10, p. 118, 2023.

[2] M. K. L. O. S. ARDEN, UMU and VERD, "Deliverable 7.2: Pilots deployment, operation and socioeconomic evaluation, h2020 phoenix," tech. rep., Horizon 2020 PHOENIX Project, 2022. European Union's Horizon 2020 Research and Innovation Programme, Grant Agreement No. 893079.

[3] J. Li, X. Li, and C. Zhang, "Analysis on security and privacy-preserving in federated learning," *Highlights in Science, Engineering and Technology*, vol. 4, pp. 349–358, 07 2022.

[4] K. Wei, J. Li, M. Ding, C. Ma, H. H. Yang, F. Farokhi, S. Jin, T. Q. S. Quek, and H. Vincent Poor, "Federated learning with differential privacy: Algorithms and performance analysis," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 3454–3469, 2020.

[5] X. Xie, C. Hu, H. Ren, and J. Deng, "A survey on vulnerability of federated learning: A learning algorithm perspective," *Neurocomputing*, vol. 573, p. 127225, 2024.

[6] N. Rodríguez-Barroso, D. Jiménez-López, M. V. Luzón, F. Herrera, and E. Martínez-Cámara, "Survey on federated learning threats: Concepts, taxonomy on attacks and defences, experimental study and challenges," *Information Fusion*, 2023.

[7] M. Moshawrab, M. Adda, A. Bouzouane, H. Ibrahim, and A. Raad, "Reviewing federated learning aggregation algorithms; strategies, contributions, limitations and future perspectives," *Electronics*, vol. 12, no. 10, 2023.

[8] W. H. Organization, "Global health estimates 2020: Deaths by cause, age, sex, by country and by region, 2000-2019," tech. rep., 2020.

[9] J. Konečný, H. B. McMahan, F. X. Yu, P. Richtárik, A. T. Suresh, and D. Bacon, "Federated learning: Strategies for improving communication efficiency," 10 2016.

[10] R. Shokri, M. Stronati, C. Song, and V. Shmatikov, "Membership inference attacks against machine learning models," in *2017 IEEE Symposium on Security and Privacy (SP)*, pp. 3–18, 2017.

[11] Z. Wang, Y. Huang, S. Mengkai, L. Wu, F. Xue, and K. Ren, "Poisoning-assisted property inference attack against federated learning," *IEEE Transactions on Dependable and Secure Computing*, vol. PP, pp. 1–13, 01 2022.

[12] C. Chen, L. Lyu, H. Yu, and G. Chen, "Practical attribute reconstruction attack against federated learning," *IEEE Transactions on Big Data*, vol. 10, no. 6, pp. 851–863, 2024.

[13] L. Itu, P. Sharma, K. Ralovich, V. Mihalef, R. Ionasec, A. Everett, R. Ringel, A. Kamen, and D. Comaniciu, "Non-invasive hemodynamic assessment of aortic coarctation: Validation with in vivo measurements," *Annals of biomedical engineering*, vol. 41, pp. 669–681, Apr. 2013. Funding Information: The authors would like to acknowledge Dr. Michael Suehling and Dr. Constantin Suciu. This work was partially supported by the Sectorial Operational Programme Human Resources Development (SOP HRD), financed from the European Social Fund and by the Romanian Government under the contract number POSDRU/88/1.5/S/76945. This work has been partially funded by European Union project Sime-Child (FP7 – 248421).

[14] A. González-Vidal, A. P. Ramallo-González, and A. Skarmeta, "Empirical study of massive set-point behavioral data: Towards a cloud-based artificial intelligence that democratizes thermostats," in *2018 IEEE International Conference on Smart Computing (SMARTCOMP)*, pp. 211–218, 2018.

[15] E. M. Campos, A. G. Vidal, J. L. Hernández Ramos, and A. Skarmeta, "Federated transfer learning for energy efficiency in smart buildings," in *IEEE INFOCOM 2023 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6, 2023.

[16] A. Zafeiropoulos, E. Fotopoulou, A. González-Vidal, and A. Skarmeta, "Detaching the design, development and execution of big data analysis processes: A case study based on energy and behavioral analytics," in *2018 Global Internet of Things Summit (GIoTS)*, pp. 1–6, 2018.

[17] A. Ntafalias, S. Tsakanikas, S. Skarvelis-Kazakos, P. Papadopoulos, A. F. Skarmeta-Gómez, A. González-Vidal, V. Tomat, A. P. Ramallo-González, R. Marin-Perez, and M. C. Vlachou, "Design and implementation of an interoperable architecture for integrating building legacy systems into scalable energy management systems," *Smart Cities*, vol. 5, no. 4, pp. 1421–1440, 2022.