UNIVERSITÉ DE BORDEAUX • MASTER 2 • 2025–2026

Swarm Learning

Critical Synthesis of “Swarm Learning: A Survey of Concepts, Applications, and Trends”

Samir Nait Sadi & Arthur Macdonald

Based on the survey by Shammar, Cui, and Al-qaness

Context & Motivation

The Problem: Data Silos

Modern AI requires massive datasets. However, data is trapped in **local silos** (hospitals, factories) due to:

- **Regulatory Barriers:**
  GDPR/HIPAA prohibit sharing raw data.
- **Centralization Risks:**
  Central servers create single points of failure.

The Objective

To enable **collaborative training** without moving raw data.

The Solution must guarantee:

- ✓ **Privacy** (Data locality)
- ✓ **Autonomy** (No central leader)
- ✓ **Security** (Trustless environment)

Roots & Evolution

Root Concept: Swarm Intelligence (SI)

Nature's solution to **Autonomy**. Biological swarms (ants/bees) solve complex tasks via local interaction, proving systems don't need a central commander.

Structural Evolution:

1. Federated (FL)

Topology: Star (Central Server)

Solved: Privacy (Silos)

Issue: Central Bottleneck

2. Decentralized (DFL)

Topology: P2P Mesh (Gossip)

Solved: Autonomy

Issue: Lack of Trust

3. Swarm (SL)

Topology: Mesh + Blockchain

Solved: Security

Result: Secure Autonomy

Comparative Analysis

Feature	Federated (FL)	Decentralized (DFL)	Swarm Learning (SL)
Topology	Star (Central Server)	Mesh (P2P Gossip)	**Mesh (P2P + Blockchain)**
Coordination	Central Aggregator	Gossip Protocols	**Smart Contracts**
Trust Model	Authority-based	Reputation-based	**Cryptographic Ledger**
Fault Tolerance	Low (Bottleneck)	Medium	**High (Byzantine Safe)**

Core Architecture

1. SL Node

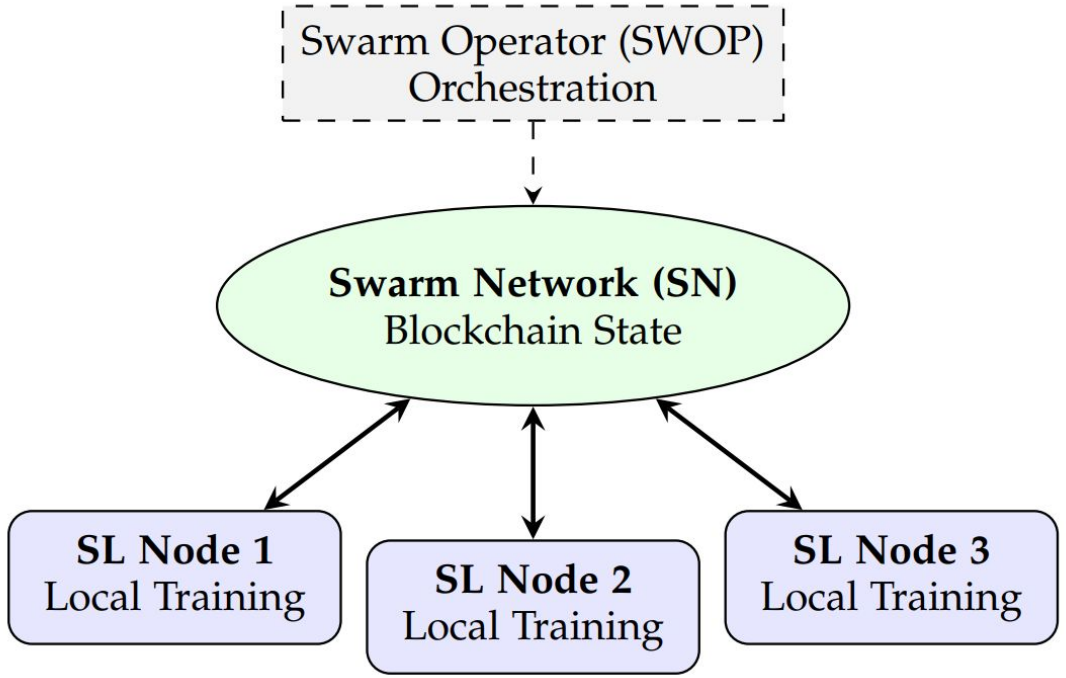
Executes local training. Sensitive data never leaves this secure boundary.

2. Swarm Network (SN)

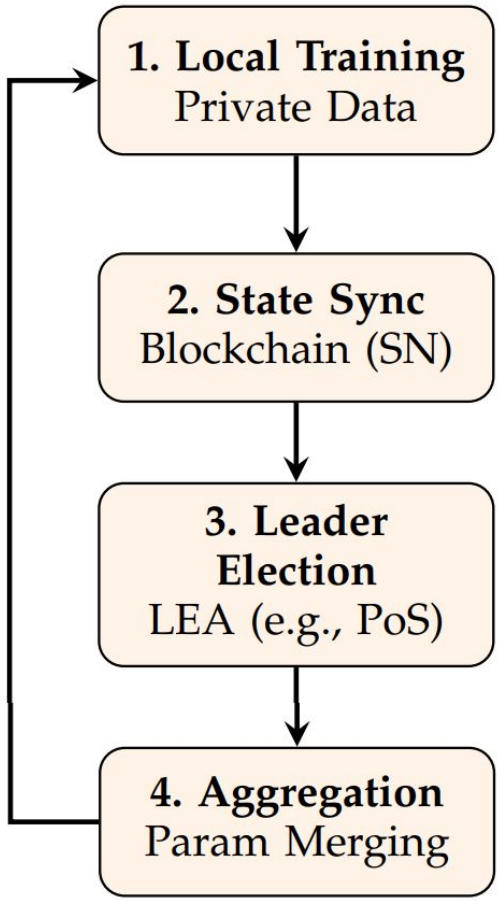
The Blockchain layer. Handles state maintenance and node registration.

3. Swarm Operator (SWOP)

Manages orchestration and identity verification (using SPIRE).



The Learning Cycle



1. Local Training

Nodes train models on private data.

2. State Sync

Metadata is shared via Blockchain.

3. Leader Election

A temporary leader is elected (e.g., PoS).

4. Aggregation

Parameters are merged into the global model.

Five Core Characteristics

Privacy

Raw data remains strictly local to the node.

Decentralization

No central authority or single failure point.

Continuous

Real-time adaptation to new data streams.

Diversity

Integrates heterogeneous datasets easily.

Real-World Applications

Healthcare

- Decentralized diagnostics (e.g., COVID-19, Leukemia).
- Strict adherence to GDPR & HIPAA regulations.

Transportation

- Coordinating autonomous vehicles via edge sensors.
- Optimizing traffic flow without central control.

Finance

- Fraud detection without exposing transaction logs.
- Secure, peer-to-peer credit scoring models.

Key Technical Challenges

Integrity Attacks

Malicious nodes may inject poisoned data or backdoors to corrupt the global model.

Privacy Leakage

Model inversion attacks can theoretically reconstruct raw data from shared gradients.

Availability Risks

Vulnerability to Eclipse attacks (node isolation) and Denial of Service (DoS).

Data Heterogeneity

Global accuracy suffers when local data distributions (Non-IID) vary significantly.

Future Research Directions

Security & Scale

- **Encryption:** Homomorphic methods for computation on encrypted data.
- **Scalability:** Optimizing protocols to handle thousands of concurrent nodes.

Efficiency & Standards

- **Green AI:** Reducing the energy cost of consensus mechanisms.
- **Standards:** Establishing interoperability for industrial adoption.

Critical Analysis of the Survey

Strengths

- **Comprehensive:** Successfully consolidates fragmented research into a unified narrative.
- **Clear Framework:** Distinguishes SL cleanly from FL and SI concepts.
- **Accessible:** Excellent entry point for researchers new to the field.

Limitations

- **Descriptive Only:** Lacks quantitative benchmarks to validate claims.
- **Practical Gap:** Minimal analysis of real-world deployment costs.
- **Vague Roadmap:** Priorities for future work are not clearly ranked.

Should You Read This Survey?

Read if you need...

- **State-of-the-Art Review:** A structured taxonomy of the SL landscape (2020-2025).
- **Conceptual Clarity:** Clear distinction between FL, DFL, and Swarm Intelligence.
- **Research Roadmap:** Identification of open gaps (Security, Non-IID) to define new projects.

Skip if you need...

- **Quantitative Benchmarks:** No direct performance metrics or accuracy comparisons.
- **Implementation Details:** Lacks code repositories, SDKs, or engineering manuals.
- **Operational Costs:** Minimal analysis of gas fees or blockchain overhead.

Conclusion

A New Standard for Collaboration SL replaces "Centralized Trust" with "Cryptographic Truth" via Blockchain, enabling secure, autonomous cooperation.

Unlocking Sensitive Domains It empowers data-siloed fields like Healthcare and Finance to leverage collective intelligence without compromising privacy (GDPR).

The Critical Path Ahead Future success hinges on solving the "Efficiency vs. Security" trade-off and establishing industrial standards.

References

[1] E. Shammar et al., "Swarm Learning: A Survey...," arXiv:2405.0055602, 2025.

[2] S. Warnat-Herresthal et al., "Swarm learning for decentralized and confidential clinical machine learning," Nature, 2021.

[3] J. Han et al., "Demystifying swarm learning: A new paradigm of blockchain-based decentralized federated learning," arXiv:2201.05286, 2022.

[4] E. T. Martínez Beltrán et al., "Decentralized federated learning: Fundamentals, state of the art, frameworks, trends, and challenges," IEEE Surveys, 2023.

[5] H. Zhang et al., "Decentralized federated learning based on blockchain: Concepts, framework, and challenges," Comp. Comm., 2024.

[6] K. Chen et al., "Backdoor attacks against distributed swarm learning," ISA Trans, 2023.

[7] H. A. Madni et al., "Blockchain-based swarm learning for the mitigation of gradient leakage in federated learning," IEEE Access, 2023.

Questions?

Thank you for your attention.