# Federated Computing

## A Data-Driven Business Infrastructure

**Enzo Fenoglio**
University College London

**Philip Treleaven**
University College London

**Abstract:** As organizations increasingly rely on distributed data assets, they require secure, scalable, and efficient infrastructures that ensure privacy, compliance, and business control. Federated Computing (FC) addresses this need as a conceptual infrastructure framework built on four key pillars: a) distributed data assets for data sovereignty, b) federated services for privacy-preserving analytics, c) standardized APIs for system interoperability, and d) distributed ledger technology for data security and collaboration. FC is modular and adaptable, allowing organizations to choose the proper modules and components that ensure seamless integration with existing platforms and regulatory requirements. The paper introduces FC's conceptual foundation, explores its deployment models, and demonstrates how it enables organizations to unlock value from distributed data assets to drive innovation, collaboration, and value creation, including both internal data utilization and external data monetization. The broader applicability of FC is pivotal for utilizing distributed data assets among external organizations and staff within an organization.

**Keywords**: Data Collaboration, Federated AI, Privacy-Preserving Computation, Data Valuation, Data Sovereignty

## 1. INTRODUCTION

The rapid expansion of data-driven technologies has intensified the demand for collaborative AI within the evolving data economy [1] [2]. As organizations become increasingly data-centric, they require federated infrastructures to utilize their distributed data assets effectively. Traditional cloud-centric models [3] often require the centralization of datasets, which introduces privacy risks, regulatory challenges, and unbalanced economic returns. Such centralization benefits platform providers more than data originators and contributors, who often receive limited recognition or compensation for the value created. This inequitable distribution of benefits becomes particularly pronounced in sectors like financial services, healthcare analytics, retail, and the industrial Internet of Things (IoT), where data sovereignty (**Section 6.2**) is critical.

Federated Computing (FC) is a comprehensive conceptual infrastructure framework enabling organizations to collaboratively analyze and derive value from distributed data assets, preserving sovereignty, ensuring regulatory compliance, and maintaining security through controlled computation at the source location. It can be considered as a meta-framework for scaffolding data-driven business infrastructure quickly. Unlike Federated Learning (FL), which is limited to training machine learning (ML) models, FC extends beyond model training to support distributed analytics, data transformations, and policy enforcement mechanisms. Most FL implementations depend on a centralized orchestrator [4], which reintroduces single points of control and limits broader data processing. FC provides a more flexible and scalable approach by integrating identity management, APIs, and configurable incentive mechanisms. FC control mechanisms can range from centralized policy enforcement to decentralized trust models, incorporating transparent ledger-based auditing and tokenized incentive structures. FC expands beyond the limitations of FL by creating a comprehensive framework for secure data collaboration across enterprise environments and cross-industry partnerships.

FC is architected around four key conceptual pillars that serve as guiding principles. These foundational elements enable organizations to design systems that align with specific business objectives, data-sharing needs, and regulatory requirements, as outlined in (**Table 1**):

1. **Distributed Data Assets**: This pillar ensures that data remains under *local ownership and control* at its source location. This pillar preserves data isolation, facilitates regulatory compliance across jurisdictions, and enhances system resilience by removing single points of failure (**Section 3.1**).

2. **Federated Services**: This pillar enables privacy-preserving analytics and model training by keeping raw data localized. It extends traditional FL for model training to include broader federated services like federated analytics and data transformation for comprehensive business applications (**Section 3.2**).

3. **Application Programming Interface (API)**: This pillar establishes standardized protocols for secure information exchange between diverse software systems. It enables consistent data exchange patterns, simplifies integration with existing enterprise infrastructures, and ensures interoperability across heterogeneous technology environments (**Section 3.3**).

4. **Distributed Ledger Technology (DLT)** [5]: This pillar leverages blockchain or similar tamper-resistant distributed record-keeping systems (ledger) to document transactions, ensure auditability, build trust, and maintain transparency. Although optional in most implementations, it can automate control mechanisms, enhance verification processes, and distribute incentives via programmable rules (smart contracts) that execute automatically when conditions are met (**Section 3.4**).

| Pillar | Operational Characteristics | Business Contribution |
|---|---|---|
| **Distributed Data Assets** | Maintains data locality with distributed storage and processing. | Reduces compliance risks and enables collaboration without data transfer. |
| **Federated Services** | A computation approach that brings algorithms to the data rather than centralizing data for processing. | Allows insights from collective data for preserving competitive information boundaries. |
| **Application Programming Interface (API)** | Provides standardized connection points with granular access control. | Enables extensible integration with existing systems and third-party services. |
| **Distributed Ledger Technology (DLT)** | Creates immutable records with cryptographic verification. | Facilitates trust between parties without requiring centralized intermediaries. |

Table 1: The Four Foundational Pillars of Federated Computing

The four pillars collectively establish a foundation for a resilient and scalable approach to federated collaboration. They address a critical challenge in today's data economy–how to extract value from distributed information. The shift is especially critical in an era where data has been described with the *new gold* or *new oil* - metaphors that underscore the complexities surrounding data's utilization. The *new gold* perspective recognizes data as a strategic asset, such as customer behavior analytics and AI model training innovations. The *new oil* emphasizes the need for controlled, structured collaboration among staff within an organization, partner organizations, social network users peer-to-peer (P2P), and, in the future, among *smart* robots and machines (cf., Industry 4.0). However, data fundamentally differs from traditional assets. Unlike physical resources, data is a *non-rivalrous asset* [6]–it can be used simultaneously by multiple parties without being consumed or diminished. This property, sometimes called *non-depletion*, distinguishes data from traditional assets like fuel or capital equipment. Data's value often increases as it is shared, reused, and recombined across contexts, particularly when insights generated in one domain can inform others. However, this same property introduces new challenges–the more data is used, the greater the exposure to privacy risks, compliance obligations, and potential misuse.
FC addresses this tension by enabling shared analytical value without centralized data movement, preserving both the utility and the sovereignty of the data creating a clear separation between several key data concepts:

- **Raw data** refers to the original information collected and stored within organizational boundaries. Raw data remains distributed under local ownership and control.

- **Data identity** represents where data resides and who controls it—preserving sovereign boundaries that ensure regulatory compliance and governance.
- **Data insights** are the analytical findings or intelligence extracted from raw data through federated computation, such as anonymized raw data, aggregated findings, or model parameters, without requiring data centralization or transfer of ownership.
- **Data Assets** are any valuable form of data insights recognized as a valuable resource. They can be managed, shared, or monetized within or across organizational boundaries.
- **Tokenized Data assets** emerge when data insights are formalized through tokenization (**Section 3.4**), creating structured, exchangeable resources. It suggests a commodity-like nature, with potential financial implications such as valuation, exchange, or even the creation of derivatives (**Section 6.5**).

This paper targets senior technical decision-makers—such as IT directors, enterprise architects, and operations leads—who are responsible for evaluating and guiding the adoption of emerging data infrastructure strategies. It offers a balanced overview of FC, providing sufficient technical insight to support informed decision-making and pilot projects guidance, without delving into the exhaustive details required for full-scale implementation or high-level strategic frameworks. With regards to deployment, FC comprises two configurations:

- **Basic FC (without DLT support)**: In this configuration, organizations derive data assets through collaborative analysis using distributed data insights, Federated Learning (FL), standardized APIs, and secure identity management systems. This approach enables the secure generation of data insights through controlled access and authentication (**Section 4.2.3**). Deployment requires configuring the necessary APIs and identity frameworks to adapt FL applications for secure analytics across distributed sources.
- **Enhanced FC (with DLT support):** This configuration extends Basic FC by integrating DLT, Decentralized Digital Identifier (DID), and tokenization. Enhanced FC transforms federated data insights into tokenized data assets that can be tracked, valued, and monetized. Additional security mechanisms, such as immutable ledger records and automated policy enforcement through smart contracts, ensure transparent and verifiable data sharing. This approach rewards high-quality data contributions and supports compliance with data protection regulations through rigorous security verification.

These deployment configurations—Basic FC and Enhanced FC—range from straightforward federated data collaboration to more complex use cases incorporating DLT and tokenization for comprehensive data valuation. The following section introduces the **Data-Token-Compute Model** to illustrate how value, data, and computation intersect within federated environments. This flexibility allows organizations to derive appropriate value from their distributed information resources, even when constrained by regulatory requirements and business objectives.

## 1.1 Data-Token-Compute Model

The Data-Token-Compute Model [7] (**Figure 1** ) establishes a mechanism for secure data collaboration and valuation by separating three critical elements: data ownership, value representation, and computational processing. This model illustrates the transformational relationships that allow organizations to extract value from distributed data resources, where compute serves as the essential bridge that transforms raw data into tokenized value without compromising ownership boundaries. By explicitly separating these domains, the model provides a conceptual structure that maintains data control yet facilitates value creation through controlled transformation processes.

**Data (ownership)**: Data remains under the full control of its original owners stored locally at its source—whether on edge devices, institutional servers, or enterprise environments. Organizations and individuals preserve privacy and comply with relevant regulations by keeping data local. Rather than transferring raw information to external processing environments, FC enables secure computational access by implementing controlled processing boundaries. This approach allows organizations to share only the resulting aggregated insights or model updates without transferring data ownership or compromising sensitive information.

**Token (value representation)**: Tokens [8] serve as cryptographically secure placeholders for the value derived from data representing ownership rights, usage permissions, provenance information, and value attribution mechanisms. Tokenization replaces sensitive data with digital tokens—secure surrogates that maintain essential utility without exposing the underlying information—creating a representation system that enables Controlled Access Rights, Transparent Reward Distribution, and Cryptographically Verifiable Transactions (**Section 3.4**). This process lays the

foundation for data marketplaces, where data value can be traded, bundled, or used to back financial instruments. Although tokenization exists outside blockchain environments (e.g., in payment processing systems like Stripe and PayPal), this paper focuses on blockchain-enabled tokenization to leverage the immutability, trust, and automated governance that DLT provides for secure token exchange and value distribution.

**Compute (processing)**: Compute is about the transformations operated on local data, such as machine learning model training, analytics processing, and any other processing tasks for data insights discovery. FC distributes computations across appropriate processing environments, such as local nodes, edge devices, or federated networks. In **Basic FC** configurations, Compute transforms raw data into data insights that directly support business applications, i.e., Raw Data (local) → (Federated Services) → Data Insights → Data Assets. In **Enhanced FC** configurations, Compute is a critical bridge between Data and Tokens, transforming raw data into tokenized data assets, i.e., Raw Data (local) → (Federated Services) → Data Insights → Tokenization (DLT) → Tokenized Data Assets. Additionally, privacy-enhancing techniques—such as differential privacy and cryptographic validation—ensure that computations remain secure, compliant, and verifiable, enabling trust across stakeholders.

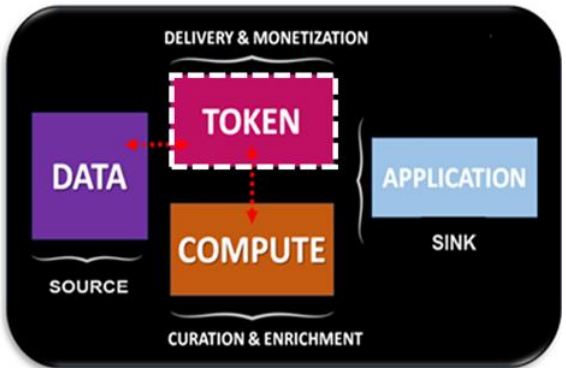


Figure 1 The Data-Token-Compute model *[7]*

(The broken line for the Token block indicates that in FC tokenization is used only with DLT)

## 2. FEDERATED SYSTEMS RELATED WORK

Federated learning (FL) [9] has emerged as a promising alternative to centralized AI. It allows AI models to be trained across distributed datasets without requiring direct access to raw information. Several frameworks have been developed to facilitate FL, each with distinct strengths and limitations. **Table 2** reviews some of the most relevant FL frameworks and highlights how FC differentiates by supporting a broader range of privacy-preserving interaction, such as, secure analytics, distributed computation, and data valuation/monetization. Key differentiators include:

| Framework | Strengths | Limitations vs FC |
|---|---|---|
| TensorFlow Federated (TFF) [10] | ✓Strong ML integration,<br>✓Scalable Federated Learning. | ✗Requires a centralized aggregation<br>✗No built-in economic incentives. |
| OpenMined PySyft [13] | ✓Privacy-preserving AI with encrypted training. | ✗Relies on trusted execution environments (TEEs);<br>✗ Lacks tokenized access control. |
| Paddle Federated Learning Framework (PaddleFL) [11] | ✓Designed for large-scale enterprise FL.<br>✓Flexible encryption option | ✗ is designed for the PaddlePaddle ecosystem<br>✗ Lacks full decentralization. |
| GAIA-X / IDSA [12] | ✓Defines governance & compliance for federated data exchange in Europe. | ✗ Lacks a cryptographic approach to privacy; focuses on policy, not FL execution. |

**Table 2** : Survey of Federated Learning frameworks compared to Federated Computing

1. **Modularity Beyond FL**: Unlike **TFF** [10] and **PaddleFL** [11], which focus primarily on federated model training, FC enables configurable data-sharing models that can operate in centralized, decentralized, or hybrid environments (see **Section 4.3**).
2. **Interoperability & Secure Collaboration**: **GAIA-X** [12] focuses on policy frameworks and compliance standards but does not implement cryptographic security measures or FL execution mechanisms. In contrast, FC integrates cryptographic techniques—such as zero-knowledge proofs, secure multiparty computation (SMPC), and smart contracts (when using DLT)—to enable verifiable, trustless transactions and secure multi-party collaboration beyond simple policy enforcement.
3. **Beyond FL Training**: Frameworks like **PySyft** [13] primarily facilitate federated model training but do not support federated analytics or secure data transformations. FC extends FL, enabling distributed services f secure data

processing, and data valuation mechanisms, making it more applicable to cross-industry collaboration and regulatory-compliant AI workflows.

4. **Configurable Access Control**: **PySyft** depends on trusted execution environments (TEEs) for privacy enforcement. On the contrary, FC provides fine-grained access control—including smart contracts, policy-based enforcement, and cryptographic guarantees. These mechanisms allow organizations to maintain ownership and control over data assets to ensure structured and compliant interactions between entities.
5. **Incentivization & Monetization**: Existing FL frameworks (e.g., **TFF, PaddleFL**) lack native economic incentives, making long-term cross-organization collaboration challenging. FC can incorporate token-based economic incentives, enabling participants to monetize insights, incentivize high-quality contributions, and maintain sustainable federated ecosystems based on specific business models (**Section 6.5**).

The remainder of the paper is organized as follows: Section 3 introduces FC's context unpacking the four pillars. Section 4 presents the modular architecture stack and illustrates how FC deployment configurations can be composed. Section 5 presents a use case for drug design and manufacture. Section 6 presents strategic business considerations for FC adoption. Section 7 discusses future challenges and concludes the paper.

## 3. FEDERATED COMPUTING CONTEXT: THE FOUR PILLARS

In today's data economy, organizations face a central tension to leverage distributed data for innovation and privacy control. This section explores the business context for Federated Computing (FC) and examines how the four pillars (**Figure 2**) address these challenges.

**The challenge of distributed data collaboration:**

Organizations view data as a strategic asset but struggle to collaborate successfully as volumes grow and become distributed across departments, locations, and external partners [14]. Traditional data-sharing models and centralized architectures introduce significant business risks:

- **Security Vulnerabilities**: Centralized repositories create single points of failure, increasing the risk of large-scale data breaches.
- **Diminished Data Sovereignty**: Organizations must relinquish direct control over sensitive information, compromising their ability to enforce data protection policies and access controls.
- **Regulatory Barriers**: International regulations restrict (raw) data movement across organizational and geographic boundaries, making centralized approaches increasingly impractical.

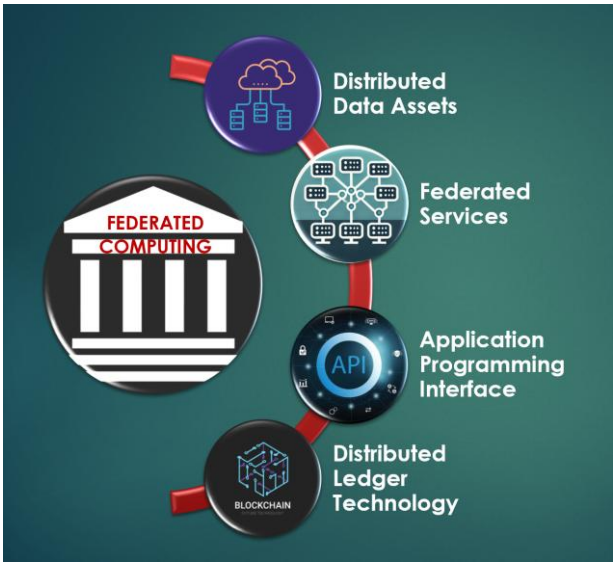


**Figure 2** Federated Computing architectural pillars

**The Federated Computing Solution:**

FC provides a flexible framework that addresses these challenges, enabling secure collaboration without centralization. Rather than consolidating data into shared repositories, FC allows computation at the data source. This architectural adaptability distinguishes FC from traditional Federated Learning (FL)—organizations can implement various operational models based on their specific business requirements and regulatory constraints, ranging from centrally orchestrated to fully distributed approaches. FC is designed to enhance data privacy by decentralizing data storage and processing. Nevertheless, it is important to acknowledge that no system can guarantee absolute privacy. Potential vulnerabilities, such as inference attacks on model updates or compromised aggregation servers, may still pose risks to sensitive information [15]. Therefore, FC should be viewed as a framework that significantly mitigates privacy risks through decentralization and modular design, recognizing the need for ongoing vigilance and implementation of complementary security measures.

The modular design of FC builds upon the foundational pillars introduced in **Section 1** and summarized in **Table 1**. These pillars can be strategically combined to create a composable architecture that enables organizations to build tailored solutions by leveraging the unique strengths of each pillar in various configurations. **Table 3** illustrates representative use cases demonstrating how different combinations of pillars facilitate distinct functional interactions within the framework. For example, an organization may adopt identity-gated execution (**UC1**), combining APIs and federated services to control access based on verified credentials without requiring blockchain infrastructure (e.g., secured API gateway for hospital analytics platforms). Another scenario may focus on privacy-preserving computation (**UC2**), where federated services interact directly with distributed data assets to run analytics without centralizing data (e.g., federated learning across diagnostic imaging centers). In other contexts, access control policies (**UC3**) might be enforced entirely through the API pillar, independent of tokenization or smart contracts (e.g., access policies in cross-border financial data sharing). Conversely, a consortium-driven environment might prioritize DLT-based policy enforcement (**UC4**) to ensure verifiability and accountability across multiple data domains (e.g., automated data-use governance in research consortia).

| | Use Cases Examples | Distributed Data Assets | Federated Services | API | Distributed Ledger Technology |
|---|---|---|---|---|---|
|  | **UC1 Identity-gated execution** | | ✓ | ✓ | |
|  | **UC2 Privacy-preserving Computation** | ✓ | ✓ | | |
|  | **UC3: Fine-grained access control** | ✓ | | ✓ | |
|  | **UC4: Smart contract-driven policy enforcement** | | | ✓ | ✓ |

**Table 3** - Pillar Participation in Federated Computing for some exemplar Use Cases.

In **Section 4,** we will demonstrate how abstract architectural pillars are implemented through concrete modules and components. But first, the following sections provide an in-depth analysis of each architectural pillar, highlighting their specific roles enabling compliant data collaboration.

## 3.1 Distributed Data Assets in FC

Modern organizations generate and collect data across multiple locations, departments, and systems. Traditional approaches often attempt to centralize into data lakes or warehouses, creating security risks and challenges. In FC, distributed data assets serve as a foundational element of the framework. Rather than consolidating data into central repositories, FC maintains data at the source, enabling secure access and computation. This approach is not unique to FC, but when combined with the other pillars, it creates a comprehensive framework for unlocking the value of distributed information resources. This infrastructure serves multiple critical functions:

- It preserves organizational data authority, empowering each department or entity to maintain authority over its data assets and determining how data are used and shared. This local control is crucial for organizations operating across regulatory jurisdictions or handling sensitive information.
- This approach naturally aligns with data isolation requirements and regulations. Organizations can implement granular access controls and maintain clear audit trails. It becomes important when dealing with personal data subject to regulations (**Section 6.4**)
- Distributed data assets enable selective sharing without compromising security. Organizations can participate in federated computations, controlling their raw data and sharing only the necessary computed results or aggregated insights. In FC, we treat all distributed data as *assets* due to its intrinsic value to the organization. This usage does not imply tokenization. Rather, it emphasizes the business importance of local data, which organizations own and control and may optionally transform into **tokenized data assets** (**Section 3.4**) for monetization.

## 3.2 Federated Services in FC

Federated learning (FL) [9] is a key computational mechanism within FC, enabling model training and privacy across distributed datasets. When integrated into FC, FL operates within a broader federated service ecosystem that introduces several enhancements beyond standard FL implementations:

- **Flexible Trust and Coordination Models**: Federated services support multiple architectural patterns for coordination and trust. The ***hub-and-spoke model*** [16] employs a central orchestrator to coordinate operations and aggregate results, providing streamlined governance for regulated environments. In contrast, the ***peer-to-peer model*** [17] distributes coordination across participating nodes without central authority, enhancing resilience and reducing central control risks. Many organizations implement hybrid approaches that combine elements of both models to align with their specific trust requirements and operational needs.
- **Configurable Data Access & Policy Enforcement**: Unlike standalone FL frameworks, FL in FC is governed by structured access policies, allowing participants to define fine-grained permissions for model updates, aggregation, and sharing.
- **Interoperability Across Federated Workflows:** Federated services in FC extend beyond model training to include federated analytics, federated queries, and secure data transformations, enabling comprehensive data collaboration pipelines adapted to specific business needs.
- **Scalability & Continuous Adaptation:** FL in FC is designed to scale across multi-party collaborations, supporting dynamic participation where new nodes can join without disrupting ongoing training. That is particularly useful in real-time federated AI applications, cross-industry partnerships, and evolving regulatory environments.

## 3.3 Application Programming Interface in FC

Application Programming Interface (APIs) in FC is the standardized communication layer that enables seamless interaction between distributed components (**Section 4**) and participants. Namely, data producers (enterprise systems, IoT devices, consumer applications); data consumers (analytics platforms, machine learning models, dashboards); model developers (creating and refining FL algorithms); regulatory authorities (requiring compliance verification and audit trails); infrastructure providers (managing computational resources and storage systems); orchestration services (coordinating operations across distributed environments). The standardized API architecture ensures that all these diverse stakeholders can interact securely and efficiently within the federated ecosystem to grant appropriate access controls.

Beyond simply enabling communication between distributed components, APIs in FC serve as the foundation for structured access control, permission management, and interoperability across diverse entities. Unlike traditional API-based integrations, which primarily focus on technical connectivity, FC APIs are designed to enforce policy-driven data exchange and streamline multi-party access control within a federated environment. APIs facilitate three essential functions:

- They provide standardized protocols for data access and computation requests. These protocols ensure that components can communicate and collaborate regardless of their proprietary technology stack or infrastructure.
- APIs enable modular integration of different federated services. Organizations can expose specific computational capabilities or data services. The modularity allows for flexible system composition and the easy addition of new services or participants.
- APIs enforce secure authentication and authorization, ensuring authorized participants access services or data insights.

The application programming interface made available to stakeholders and software developers supports APP cooperation. APIs typically operate through data exchange standards like JSON or XML. Examples include:

- **JSON API** is an application programming interface for lightweight data interchange (text-based data exchange format) between computer applications.
- **REST API** also called a RESTful API or RESTful web API, conforms to the design principles of the representational state transfer (REST) architectural style. It uses HTTP requests to access and use data.
- **Open API** defines a standard, language-agnostic interface to data and APIs that supports data access (e.g., JSON, XML, CSV) and cooperation of applications and allows applications to be developed independently by third parties.

### 3.4 Distributed Ledger Technology in FC

DLT is the fourth pillar and provides an additional layer of trust and transparency that enhances secure business operations within FC when required. A distributed ledger is a synchronized database architecture maintained across multiple organizations, eliminating reliance on a single authority for data validation and integrity. Blockchain [18] technology is the most widely adopted form of DLT, organizing data into cryptographically linked blocks that form an immutable chain of records [19]. Alternative non-blockchain implementations, such as Hedera Hashgraph [20], employ a directed acyclic graph (DAG) consensus mechanism to achieve higher transaction speeds with improved energy efficiency. These architectures ensure that once transactions are recorded, they cannot be altered without detection, providing a robust foundation for business operations requiring high trust, transparency, and auditability. Key components of blockchain include **timestamps** (record the precise time and date of block creation); **hash values** (serve as unique digital identifiers, like fingerprints, that secure block information); **nonces** (one-time-use values that enhance security by ensuring cryptographic hash algorithms and authentication protocols resist tampering). When integrated with the other three pillars of FC, DLT enables advanced capabilities through technologies like smart contracts—self-executing agreements that automatically enforce predefined rules. These infrastructure technologies strengthen FC implementations that require transparent record-keeping and incentive distribution mechanisms [5]. In the context of FC, DLT covers three critical business functions:

- It provides an immutable audit trail of all federated operations. Key federated operations, such as data access requests and computation execution, can be recorded in the ledger with cryptographic verification. It creates a transparent and verifiable history for auditing requirements.

- It enables automated governance through smart contract-based control of federated interactions. **Smart contracts** enforce predefined business rules and conditions without intermediary intervention by reducing operational overhead, eliminating potential disputes over contract execution, and ensuring consistent enforcement of business agreements across participants. When organizations share computed insights from their data, these contracts automatically enforce access controls and compensation agreements.

- It facilitates the tokenization of data value to become **tokenized data assets**, where tokenization [8] creates a digital representation of a real thing, such as a data asset, enabling measurable value attribution and controlled exchange within federated ecosystems. This transformation process introduces three key capabilities: **Controlled Access Rights** (only authorized parties can redeem or interact with tokens' underlying value as defined by smart contracts or access policies), **Transparent Reward Distribution** (smart contracts automate incentive mechanisms, distributing compensation fairly among data contributors or stakeholders), **Cryptographically Verifiable Transactions** (token exchanges and value transfers are permanently recorded and traceable, enabling organizations to measure their data contributions to federated analyses and receive fair compensation through automated distributions.)

## 4. FEDERATED COMPUTING ARCHITECTURAL STACK

Federated Computing (FC) Federated Computing (FC) implements a structured framework for extracting value from distributed data. Before exploring specific implementation components, it is important to understand the relationship between FC's conceptual foundation and its operational structure.

FC's architecture consists of two key structural elements: **conceptual pillars** defining the architectural principles (**Section 3**) and **operational modules** implementing these principles. This distinction allows organizations to maintain architectural integrity and adapt implementations to specific contexts.

### 4.1 Conceptual Pillars vs Operational Modules.

**Section 3** introduced the four pillars of FC as foundational architectural elements that guide its structure, behavior, and trust model. We also demonstrated how these pillars can be combined to address specific use cases (**Table 3**). To understand the relationship between these conceptual pillars and the operational modules described in this section, consider an object-oriented programming analogy where pillars function as abstract classes that define architectural roles without specifying implementation details. In contrast, modules serve as concrete classes that inherit and implement these roles in specific operational contexts. Thus, pillars represent abstract responsibilities that shape the design of collaboration systems, not strict implementation requirements. Each module implements one or more roles defined by the corresponding pillars. For example, the Federated Processing Backend module manages distributed execution to enforce data locality and implement aspects of the Federated Services and

Distributed Data Assets pillars. This approach creates a flexible, composable architecture that organizations can adapt to their priorities and contexts. It maintains a clear separation between conceptual design and implementation.

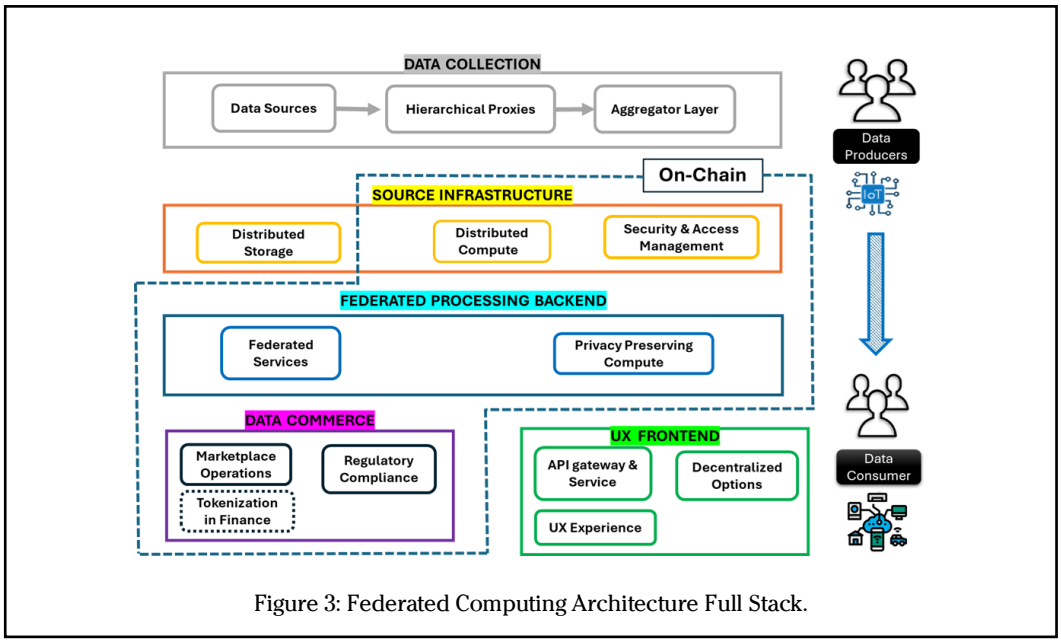


Figure 3: Federated Computing Architecture Full Stack.

## 4.2 Operational Modules Overview

The FC architecture consists of five interconnected operational modules (**Figure 3**), each translating the conceptual pillars into practical implementation components. These modules form a comprehensive transformation pathway from raw data to business value, creating a flexible architecture that organizations can adapt to their needs.

Unlike traditional data infrastructure that requires standardized implementation, FC's modular design allows organizations to select and configure components based on their unique requirements, existing systems, and regulatory constraints. This adaptability enables incremental adoption, where organizations can begin with core capabilities and expand as their federated ecosystem matures. The following sections provide an overview of each module, highlighting their business functions, relationships with the conceptual pillars, and real-world applications.

### 4.2.1 Data Collection

The Data Collection module acquires, validates, and securely processes raw data before it enters the federated ecosystem. Unlike centralized pipelines that concentrate data in one location, FC's distributed acquisition model maintains data locality, enhancing system resilience through redundancy across independent sources. This module implements aspects of the Distributed Data Assets pillar by ensuring that source data remains under local control throughout the collection process. It establishes the foundation for data authority by performing preliminary validation at the source, eliminating problematic data before it enters the federated ecosystem. Key functions include:

- **Capturing raw data from various producers**: Whether from enterprise systems, IoT devices, or consumer applications, the module provides standardized interfaces that accommodate diverse data sources for preserving their original context and ownership
- **Validating inputs to filter incomplete or malformed data**: By applying business rules and data quality checks at the edge, the module ensures that only high-quality data enters the federated environment, reducing downstream processing errors.
- **Ensuring data integrity across federated environments**: The module maintains cryptographic verification and consistent format validation to preserve data authenticity throughout the federated ecosystem.

**For example**, in a healthcare collaboration scenario, this module enables multiple hospitals to validate patient data according to their specific privacy policies before making anonymized insights available for collaborative research without centralizing protected health information.

Organizations can implement this module with varying levels of complexity depending on their scale requirements, from simple load-balancing approaches for smaller deployments to sophisticated hierarchical architectures for enterprise-scale implementations.

#### 4.2.2 Source Infrastructure

The Source Infrastructure module is the secure foundation for FC's data independence approach, managing the storage, computation, and access control resources supporting distributed data assets. Rather than simply providing storage capabilities, this module creates a comprehensive environment where data can be securely maintained, processed, and shared according to organizational policies. This module directly implements the Distributed Data Assets pillar, establishing the necessary infrastructure for the Federated Services pillar. Preserving data at its source and enabling controlled access resolves the fundamental tension between data utilization and protection. It consists of three key components:

- **Distributed Storage**: Enable organizations to maintain data within their existing infrastructure boundaries, making it accessible for federated processing. This storage layer adapts to various business requirements, supporting everything from distributed networks for high redundancy to traditional enterprise databases for performance-critical applications.
- **Distributed Compute**: Facilitate preliminary data processing at the source before broader federated analytics take place. This localized computation reduces network traffic, enhances security by minimizing data movement, and allows organizations to apply proprietary transformations to maintain complete control over sensitive information.
- **Security & Access Management:** Create a structured framework for controlled data access based on verifiable credentials and explicit permissions. This component ensures that only authorized entities can access specific data elements under clearly defined conditions, maintaining regulatory compliance and organizational data governance.

**For example**, in a financial services scenario, this module allows banks to maintain customer transaction data within their secure environments and enable anti-fraud analysis across institutions. The distributed computation capabilities enable the detection of suspicious patterns without exposing individual customer records. The security layer ensures that only authorized queries conforming to regulatory requirements can access the data.

Organizations can participate in broader federated ecosystems without compromising their data compliance obligations by creating this secure infrastructure at the data source. The flexibility of this module allows it to adapt to various organizational needs, from highly regulated environments to collaborative research initiatives

#### 4.2.3 Federated Processing Backend

The Federated Processing Backend module serves as the computational core of the FC framework, enabling secure collaborative analysis across organizational boundaries without centralizing sensitive data. This module transforms the traditional paradigm of "bringing data to computation" into "bringing computation to data," preserving privacy and unlocking collaborative insights. It directly implements the Federated Services pillar, respecting the constraints established by the Distributed Data Assets pillar. It represents the most significant innovation within the FC framework by enabling cross-organizational analytics. It consists of two complementary components:

- **Federated Services**: Provide distributed execution environments that orchestrate collaborative computation across multiple data sources. These services coordinate model training, analysis tasks, and insight generation, respecting each organization's data boundaries. Whether implementing machine learning models that learn from distributed datasets or conducting complex analytics across heterogeneous sources, these services ensure that raw data remains at its source and that derived insights can be securely shared.
- **Privacy-Preserving Compute**: Employ advanced techniques to ensure data confidentiality even during processing. From secure hardware environments to cryptographic methods, these approaches provide mathematical guarantees that computation can occur without exposing underlying data. This component enables trusted execution in potentially untrusted environments, expanding the scope of possible collaborations.

**For example**, in pharmaceutical research, this module allows multiple hospitals to collectively train disease prediction models across diverse patient populations without sharing protected health information. Each institution runs computations locally on patient data, sharing only encrypted model updates rather than sensitive records to ensure no individual patient data can be reverse engineered from the shared insights.

This balance between computational power and data protection creates a foundation for previously impossible collaborations in highly sensitive domains like healthcare, finance, and government. Organizations can extract collective intelligence from distributed data assets while maintaining complete control over their raw information.

#### 4.2.4 Data Commerce

The Data Commerce module establishes the infrastructure for quantifying, exchanging, and monetizing the value derived from federated data insights. By creating structured data valuation and exchange mechanisms, this module transforms abstract data utility into concrete business assets that can be managed within established financial and regulatory frameworks [21]. This module implements aspects of the Distributed Ledger Technology pillar (when using blockchain-based approaches) and extends the value proposition of the Federated Services pillar by creating sustainable economic incentives for collaboration. It addresses a critical challenge in multi-party data ecosystems: ensuring fair compensation for data contributions without exposing the underlying information. It consists of three integrated components:

- **Tokenization in Finance**: Create tradeable representations of data insights that can be exchanged without revealing source data. This component establishes the foundation for data financialization by transforming data values into fungible or non-fungible tokens that represent rights to insights, analysis results, or model outcomes rather than the data itself.
- **Marketplace Operation**: Provide the governance, transaction, and settlement infrastructure for secure data asset exchange. This component establishes rules for fair pricing, automated policy enforcement, and transparent value distribution among contributors. By creating standardized exchange mechanisms, it reduces friction in data markets and ensures that value attribution aligns with actual contribution
- **Regulatory Compliance**: Ensure data exchanges adhere to data protection regulations and financial oversight requirements. This component bridges the gap between data governance and financial compliance, making data assets compatible with institutional investment frameworks and regulatory obligations.

**For example**, in retail analytics, this module transforms consumer spending trends from multiple retailers into tradeable derivative instruments on digital marketplaces. Brands and investors can purchase these insights or hedge against market behavior without accessing individual store data. The tokenization mechanism creates financial assets while the distribution system automatically compensates contributing retailers based on their data's value, ensuring fair compensation without exposing sensitive sales records.

#### 4.2.5 UX Frontend

The UX Frontend module is the critical interface layer between the FC infrastructure and its diverse users, making complex federated capabilities accessible through intuitive, purpose-built experiences. By abstracting the underlying technical complexity, this module enables organizations to integrate federated computing into their existing workflows and applications without extensive retraining or disruption. This module primarily implements the Application Programming Interface (API) pillar, creating standardized interaction patterns that shield users from the intricacies of distributed systems, providing controlled access to federated capabilities. It bridges the technical sophistication of the FC framework with the practical usability needs of various stakeholders. It consists of three integrated components:

- **API Gateway & Service**: Provide a unified access layer for authentication, request routing, and response management across the federated ecosystem. This component establishes consistent communication protocols, enabling seamless integration with enterprise applications, analytics dashboards, and business intelligence tools. Creating standardized API patterns allows organizations to leverage federated capabilities within familiar technical environments.
- **Decentralized Options**: Extend the frontend capabilities to support blockchain interoperability and decentralized applications when required. This component enables organizations implementing the full FC stack with DLT to interact with on-chain governance, tokenized assets, and distributed identity systems through conventional application interfaces, reducing the technical barriers to advanced decentralized capabilities.
- **UX Experience**: Focus on developer productivity and user engagement through intuitive interfaces tailored to specific business contexts. This component ensures that the power of federated computing remains accessible to both technical and non-technical users through appropriate abstractions, visualizations, and workflow integrations.

**For example**, in a cross-border financial compliance scenario, this module allows risk analysts to query federated transaction datasets through familiar dashboard interfaces without understanding the underlying cryptographic protocols that protect sensitive banking information. The API layer translates their compliance queries into appropriate federated operations. The UX components visualize results in context-relevant formats that highlight potential regulatory issues.

This approach reduces FC adoption barriers by integrating powerful federated capabilities into existing business workflows and technical environments. Organizations can leverage the full potential of FC without requiring extensive retraining or technical expertise across their workforce.

#### 4.2.6 DLT Integration Implications

When organizations implement the Distributed Ledger Technology pillar, several components across the FC modules transition to on-chain deployment, as illustrated in **Figure 3** for the FC Enhanced configuration **(Section 1)**. This architectural choice fundamentally alters module interaction and introduces new capabilities and considerations.
The modules affected by DLT integration include:

- **Source Infrastructure** – The Security & Access Management component transitions to blockchain-based identity and permission systems, enabling cryptographically verifiable access control that operates independently of centralized authorities.
- **Federated Processing Backend** – Storage remains off-chain, but coordination and orchestration aspects transition to on-chain implementation, enabling decentralized federated learning and analytics with transparent governance. This hybrid approach maintains data locality, leveraging blockchain for verifiable computation tracking.
- **Data Commerce** – This module becomes fully on-chain, with tokenization, marketplace operations, and regulatory compliance mechanisms implemented through smart contracts that provide automated, transparent, and tamper-resistant governance.
- **UX Frontend** – This module remains primarily off-chain but is enhanced with capabilities to interface directly with on-chain federated services, creating a seamless user experience regardless of the underlying implementation. This approach shields users from blockchain complexity enabling them to benefit from its security and transparency features.

DLT integration enables enhanced trust mechanisms through immutable audit trails, automated policy enforcement via smart contracts, and transparent value attribution for data contributors. These capabilities are particularly valuable in scenarios requiring collaboration between organizations with limited prior relationships or regulatory environments demanding provable compliance. However, this integration also introduces additional complexity, performance considerations, and governance requirements that organizations must carefully evaluate against their specific business needs. Moreover, determining which data should be stored on-chain versus off-chain is crucial to balance transparency with performance and privacy [22]. The modular nature of FC allows organizations to selectively implement DLT components only where they provide clear business value rather than mandating blockchain adoption across the entire architecture. This flexibility enables tailored solutions that align with specific organizational needs and constraints.

## 5. USE CASE: DRUG DESIGN AND MANUFACTURE (PANDEMIC RESPONSE)

This section illustrates how Federated Computing (FC) supports accelerated pharmaceutical innovation in pandemic scenarios by enabling secure, compliant collaboration across stakeholders with sensitive data. It highlights how FC modules (**Section 4**) can be composed to address industry-specific challenges—ensuring privacy, regulatory adherence, and protection of proprietary information—demonstrating practical deployment guidance in a highly regulated environment.

**The Problem:**

During a pandemic crisis (e.g., COVID-19), multiple stakeholders—national health services, research institutions, AI analytics companies, pharmaceutical manufacturers, and national biobank databases—must collaborate to accelerate drug discovery and production, ensuring data privacy and security. However, direct data sharing is restricted due to data regulations that limit access to sensitive patient data and proprietary research concerns, as organizations need to protect their intellectual property and competitive restrictions, where pharmaceutical companies hesitate to share data with competitors.

**The Solution:**

FC enables secure collaboration, allowing stakeholders to pool insights without sharing raw data. In this model:

- Hospitals and national biobanks retain full control over patient records.
- Pharmaceutical and AI firms train drug discovery models using Federated Learning (FL) and secure computation environments, reducing data-sharing risks, accelerating time-to-market for life-saving drugs.

- Trusted Execution Environments (TEEs) ensure that AI simulations run across multiple datasets to preserve data confidentiality.
- Pharmaceutical companies can evaluate drug efficacy on diverse patient cohorts without direct access to sensitive records.

**Access Control and Compliance:**
Data privacy and regulation are becoming increasingly important:

- Access control is managed through enterprise Identity and Access Management (IAM) models to securely regulate participation for research institutions and regulatory bodies (**Section 6.3**).
- Regulatory compliance mechanisms provide real-time auditability, supporting fast-track approvals from agencies like the FDA and EMA.
- Research institutions retain ownership of AI-generated insights, which can be licensed to pharmaceutical firms through traditional agreements—without requiring tokenization or blockchain transactions.

**Impact, Accelerating Drug Development Securely:**
Federated Processing and Secure Infrastructure enhances drug discovery by achieving two critical objectives simultaneously:

- Preserving complete data control, enabling robust multi-party scientific collaboration.
- Maintaining clear business incentives by allowing research institutions to monetize their insights without compromising proprietary information.

This modular configuration provides a scalable and practical solution for the pharmaceutical industry. It enables federated AI-driven innovation without requiring blockchain support or tokenized data exchanges.

## 6. STRATEGIC BUSINESS CONSIDERATIONS FOR FEDERATED COMPUTING ADOPTION

In today's landscape, a substantial portion of data insights remains locked within organizational silos [14] [23], limiting their potential impact. Regulatory constraints, and a general lack of trust in data-sharing mechanisms prevent organizations from fully utilizing these datasets—leading to underutilized, duplicated, or fragmented information scattered across isolated infrastructures [24]. FC encourages secure cross-industry collaboration and innovation by strictly separating data (which stays within organizational boundaries) from its value (which can be measured and rewarded). This section examines key decision factors for senior managers and executives, including collaboration models that align with business objectives that satisfy governance requirements, interoperability standards that reduce integration costs, and data valuation approaches that create sustainable business value.

### 6.1 Federated Collaboration

Federated Computing creates a foundation for diverse collaborative models, transforming how organizations and individuals leverage distributed data assets. This infrastructure supports industry interaction patterns [25], each addressing specific business needs.

**Business-to-Business (B2B) Collaboration**: Federated Computing enables cross-industry data sharing and analytics within strict organizational boundaries. Organizations can combine complementary datasets to generate enhanced insights without centralizing sensitive information.
**Business-to-Consumer (B2C) Engagement**: FC transforms the relationship between organizations and individuals by enabling personalized services that preserve consumer privacy, where individual data contributions can be valued and rewarded through transparent mechanisms.
**Consumer-to-Consumer (C2C) Data Cooperation:** A particularly noteworthy application of FC is its ability to enable direct collaboration between individuals through structured data cooperatives and peer-to-peer marketplaces.

These collaboration models are supported through several key mechanisms:

- **Collaborative Value Creation**: Organizations or individuals pool their data analytics insights to create aggregated value without sharing raw data.
- **Cross-Domain Collaboration** [26]: FC supports **B2B**, **B2C**, and **C2C** collaboration models.

- **Data Cooperatives (Data Co-Ops)** [27]: Individuals or organizations form collaborative structures to pool and collectively benefit from shared data. These member-owned entities allow participants to aggregate their data power, implement democratic control over data usage, negotiate favorable terms with data consumers, and ensure ethical data utilization. Within FC, data cooperatives can maintain distributed raw data storage, enabling federated analytics and automated value distribution according to collectively established rules.

## 6.2 Data Sovereignty and Policy-Driven Data Sharing

Organizations face increasing pressure to derive value from their data and preserve control, security, and privacy. Traditional data-sharing models often require organizations to trust centralized intermediaries, exposing them to data breaches, and regulatory challenges. FC introduces *policy-driven data sharing* that preserves organizational sovereignty, allowing entities to control their digital identities without depending on centralized authorities. Key Aspects of Policy-Driven Data Sharing in FC:

- **Computation at the Source**: Rather than requiring data migration to external repositories, FC brings analytics and computational processes to where data resides. This approach allows organizations to maintain data within their existing infrastructures, whether on-premises, hybrid environments, or preferred cloud platforms (e.g., AWS, Azure, GCP).

- **Controlled, Non-Custodial Access**: Organizations maintain complete authority over how their data is accessed and utilized. The FC approach ensures that only explicitly authorized computations are performed on distributed data, with no raw data transfers required. Organizations define specific access rules and computational boundaries, enabling external partners to extract valuable insights without exposing underlying datasets.

- **Selective Information Exchange**: Instead of sharing complete datasets, FC enables highly targeted information sharing based on organizational policies. Organizations can selectively release **cryptographic proofs** that verify specific data properties (e.g., proving a dataset meets compliance requirements), **anonymized insights** (e.g., summary statistics on patient outcomes across hospitals), or **aggregated results** (e.g., regional energy consumption patterns without disclosing individual meter data). This granular approach minimizes exposure risk and maximizes collaborative value.

- **Data Protection by Design**: The policy-driven approach aligns with global data protection frameworks and industry-specific regulations. FC ensures that data residency requirements are respected by keeping data in its original location. The system maintains comprehensive audit trails of all computations and access patterns, enforcing granular controls at the policy level to ensure regulatory alignment.

## 6.3 Interoperability and Standardization

Ensuring seamless integration across systems is critical for adopting FC in enterprise environments. Interoperability allows organizations to enable a data business without requiring disruptive changes to their existing infrastructure. FC emphasizes standardization at multiple levels to ensure compatibility between different technologies and regulatory frameworks:

- **Cross-Platform Compatibility**: FC can be integrated with on-premises systems, cloud environments (AWS, Azure, GCP), and hybrid infrastructures. Standardized APIs and protocol-based communication ensure organizations can interact with FC using their existing IT ecosystems.

- **Standardized APIs and Protocols**: FC promotes open, industry-standard APIs (e.g., REST, gRPC, GraphQL) and communication protocols to enable smooth data exchange between different platforms.

- **Cross-Compliance Mechanisms**: Different jurisdictions have varying data protection laws, making regulatory verification challenging in federated ecosystems. FC ensures interoperability by allowing organizations to define policy-driven data-sharing rules that align with their regulatory requirements.

- **Hybrid Data Processing Models**: Some organizations may prefer centralized processing (for regulatory reasons), and others may operate in decentralized environments. FC enables hybrid models, where organizations can mix local processing, federated AI training, and cloud-based analytics within the same ecosystem.

## 6.4 Configurable Access Control and Trust Mechanisms

Effective access control over data and usage is essential for organizations operating in federated environments. Traditional data-sharing models typically rely on centralized intermediaries to enforce policies and manage access, often creating bottlenecks and potential points of failure. FC addresses this challenge by providing configurable access control models that align with specific business requirements, risk tolerance, and regulatory obligations. Organizations can choose between centralized, hybrid, or enterprise-controlled mechanisms based on their operational needs. Key Control & Trust Mechanisms in FC:

- **Tiered Access Control Models:** FC supports multiple frameworks that can be configured based on organizational requirements. These include Identity and Access Management (IAM) and Decentralized Identity (**DID**) systems for cross-organizational authentication without centralized identity providers (**Section 4.2.3**).

- **Confidential Computing & Secure Execution**: FC protects information even during computation through trusted execution environments (TEEs) such as Intel **SGX, AMD SEV,** and **AWS Nitro Enclaves**. These technologies create isolated, secure processing areas within hardware, whilst complementary cryptographic techniques provide additional protection for specific use cases.

- **Verification & Auditable Operations**: FC provides comprehensive auditability through traditional logging and optional immutable records. The system maintains detailed audit trails for all data access, computation requests, and cross-organizational transactions. Organizations can optionally implement distributed ledger technology for tamper-evident audit records in highly regulated environments and automated verification through policy-as-code frameworks.

This multilayered approach enables organizations to implement precisely the level of control required for their specific business context. FC adapts to diverse regulatory and business requirements, allowing participants to maintain their policies without compromising interoperability or security.

## 6.5 Data Valuation and Incentive Mechanisms

Accurately valuing data and creating appropriate incentive structures is critical to sustainable data collaboration [7]. Within the Federated Computing framework, organizations can transform abstract data utility into measurable business value, ensuring fair compensation for human or machine participants (**Section 3.3**).

**Understanding Data Value Creation**: Data valuation and monetization are complementary but distinct functions in the data value chain:

- **Data Valuation:** An internal, engineering-driven process that assesses the intrinsic worth of data by evaluating its quality, uniqueness, and potential utility [28]. Organizations typically leverage algorithmic valuation models, metadata analytics, and benchmarking methods to quantify data's relative importance and inform strategic decisions regarding usage, protection, and investment.

- **Data Monetization:** An external, business-driven process that converts data value into measurable financial returns or competitive advantages [29]. Monetization includes strategies such as developing data products and insights-as-a-service offerings, creating secure data marketplaces and exchanges, establishing strategic data partnerships to combine complementary data sets, implementing value-based attribution models to distribute revenues fairly, and enhancing products or services with proprietary insights to deliver differentiated customer experiences.

**Federated Computing's Contribution to Data Valuation**: Traditional data monetization often relies on centralized platforms, where data contributors have limited control over usage and revenue distribution. FC introduces a modular approach to monetization, allowing organizations to select appropriate valuation and compensation mechanisms based on business needs and regulatory constraints (**Section 4.4**) according to these principles:

- **Transparent Value Attribution**: FC implements mechanisms that track data contributions throughout the analytics lifecycle, ensuring appropriate credit and compensation for participants. This approach extends beyond simple data provision to include model improvement, feature engineering, and domain expertise.

- **Configurable Incentive Frameworks**: Organizations can implement multiple incentive models, such as *reputation-based systems* that recognize consistent, high-quality data contributions*; Access reciprocity*

*models* where contributors gain equivalent access to collaborative insights; *Sliding-scale compensation* structures that reward contributions proportional to their impact; *Tiered participation frameworks* that accommodate different levels of data sharing comfort.

- **Value-Preserving Analytics**: FC preserves local data's strategic value, enabling collaborative value creation and sharing only derived insights. This approach allows organizations to realize the benefits of the data ecosystems.

- **Auditability and Verification**: FC's access control mechanisms provide transparent records of how data is used and what value it generates, creating the foundation for fair compensation. This approach is essential for organizations that must demonstrate return on investments.

These strategic considerations—from collaboration models to data valuation mechanisms—collectively provide a framework for organizations to evaluate and implement FC in ways that align with their business objectives. The following section concludes our discussion by summarizing FC's contributions and identifying key areas for future research.

## 7. CONCLUSION

Federated Computing (FC) provides a comprehensive meta-framework that enables secure, privacy-preserving, and scalable data collaboration across organizational boundaries. Unlike traditional Federated Learning, FC extends beyond model training to include federated analytics, structured data valuation, and policy enforcement mechanisms. Built on four foundational pillars—Distributed Data Assets, Federated Services, Standardized APIs, and optional Distributed Ledger Technology—FC offers organizations a flexible approach to achieving data control while complying with regulatory requirements.

This framework addresses critical business challenges in today's data economy by maintaining data ownership at its source while enabling collaborative value creation. The modular architecture allows organizations to implement configurations ranging from minimalist deployments focusing on core federated services to fully decentralized systems with tokenized data assets.

Future research should focus on optimizing data isolation computation to balance security with performance, enhancing access control automation, expanding cross-platform interoperability, and developing standardized evaluation metrics. These advancements will be crucial for widespread FC adoption, particularly in regulated industries such as healthcare, finance, and supply chain management

## AUTHORS

**Enzo Fenoglio** is a CS Hon. Sr. Research Associate at University College London (UCL) and former AI technical lead at Cisco Systems. He specializes in DL methods for IoT platforms, DLT for Web 3.0 applications, and AI ethics. Enzo is an accomplished international speaker at AI events with 40+ patents in ML/DL. Contact him at e.fenoglio@ucl.ac.uk

**Philip Treleaven** is Professor of Computer Science at University College London and director at the UK Centre for Financial Computing & Analytics. His research interests include data science, algorithms, and blockchain technologies. He received a Ph.D. from The University of Manchester. He is a Member of the IEEE and the IEEE Computer Society. Contact him at p.treleaven@ucl.ac.uk.

## REFERENCES

[1] M. Mariniello, "The Data Economy," in *Digital Economic Policy: The Economics of Digital Markets from a European Union Perspective*, Oxford University Press, 2022.

[2] O. Iglesias, M. and A. González-Agote, "The Future Data Economy —Competitive, Fair, Safe,," IE CGC, 2024.

[3] W. Voorsluys, J. Broberg and R. Buyya, "Introduction to Cloud Computing," *Cloud Computing,* p. 1–41, January 2011.

[4] I. Kholod, E. Yanaki, D. Fomichev, E. Shalugin, E. Novikova, E. Filippov and M. Nordlund, "Open-Source Federated Learning Frameworks for IoT: A Comparative Review and Analysis," *Sensors,* vol. 21, p. 167, December 2020.

[5] R. Soltani, M. Zaman, R. Joshi and S. Sampalli, "Distributed Ledger Technologies and Their Applications: A Review," *Applied Sciences,* vol. 12, 2022.

[6] C. I. Jones and C. Tonetti, "Nonrivalry and the Economics of Data," *American Economic Review,* vol. 110, p. 2819–58, September 2020.

[7] H. Pithadia, E. Fenoglio, B. Batrinca, P. Treleaven, R. Echim, A. Bubutanu and C. Kerrigan, "Data Assets: Tokenization and Valuation," *SSRN Electronic Journal,* 2023.

[8] M. di Angelo and G. Salzer, "Tokens, Types, and Standards: Identification and Utilization in Ethereum," in *2020 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)*, 2020.

[9] P. Treleaven, M. Smietanka and H. Pithadia, "Federated Learning: The Pioneering Distributed Machine Learning and Privacy-Preserving Data Technology," *Computer,* vol. 55, pp. 20-29, 2022.

[10] TensorFlow_Federated, "TensorFlow Federated: Machine Learning on Decentralized Data," [Online]. Available: https://www.tensorflow.org/federated.

[11] Paddle_FL, "An Open-Source Deep Learning Platform Originated from Industrial Practice," 2024. [Online]. Available: https://www.paddlepaddle.org.cn/en.

[12] A. Braud, G. Fromentoux, B. Radier and O. Le Grand, "The Road to European Digital Sovereignty with Gaia-X and IDSA," *IEEE Network,* vol. 35, no. 2, pp. 4-5, 2021.

[13] A. Ziller, A. Trask, A. Lopardo, B. Szymkow, B. Wagner, E. Bluemke, J.-M. Nounahon, J. Passerat-Palmbach, K. Prakash, N. Rose, T. Ryffel, Z. N. Reza and G. Kaissis, "PySyft: A Library for Easy Federated Learning," 2021, pp. 111-139.

[14] J. Li, B. Li, X. Liu, R. Xu, J. Ma and H. Yu, "Breaking Data Silos: Cross-Domain Learning for Multi-Agent Perception from Independent Private Sources," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[15] J. Zhao, S. Bagchi, S. Avestimehr, K. Chan, S. Chaterji, D. Dimitriadis, J. Li, N. Li, A. Nourian and H. Roth, "The Federation Strikes Back: A Survey of Federated Learning Privacy Attacks, Defenses, Applications, and Policy Landscape," *ACM Computing Surveys,* vol. 57, p. 1–37, April 2025.

[16] F. Calefato and F. Lanubile, "A Hub-and-Spoke Model for Tool Integration in Distributed Development," in *2016 IEEE 11th International Conference on Global Software Engineering (ICGSE)*, 2016.

[17] C. Bussler, "P2P in B2BI," in *Proceedings of the 35th Annual Hawaii International Conference on System Sciences*, 2002.

[18] K. Wüst and A. Gervais, "Do you Need a Blockchain?," in *2018 Crypto Valley Conference on Blockchain Technology (CVCBT)*, 2018.

[19] M. N. M. Bhutta, A. A. Khwaja, A. Nadeem, H. F. Ahmad, M. K. Khan, M. A. Hanif, H. Song, M. Alshamari and Y. Cao, "A Survey on Blockchain Technology: Evolution, Architecture and Security," *IEEE Access,* vol. 9, p. 61048–61073, 2021.

[20] D.-M. Alahmad, I. Alshaikhli, A. Alkandari, A. Alshehab, M. Islam and M. Alnasheet, "Influence of Hedera Hashgraph over Blockchain," *Journal of Engineering Science and Technology,* vol. 17, pp. 3475-3488, October 2022.

[21] E. Napoletano and J. Schmidt, Decentralized Finance Is Building A New Financial System, www.forbes.

[22] F. Javed, E. Zeydan, J. Mangues-Bafalluy, K. Dev and L. Blanco, *Blockchain for Federated Learning in the Internet of Things: Trustworthy Adaptation, Standards, and the Road Ahead,* 2025.

[23] J. R. Kancharla and S. D. Madhu Kumar, "Breaking Down Data Silos: Data Mesh to Achieve Effective Aggregation in Data Localization," in *2023 International Conference on Computer, Electronics & Electrical Engineering & their Applications (IC2E3)*, 2023.

[24] S. James and A. D. Duncan, "Over 100 Data and Analytics Predictions Through 2028," Gartner Research, 23 April 2023. [Online]. Available: https://www.gartner.com/en.

[25] J. Buford, K. Mahajan and V. Krishnaswamy, "Federated enterprise and cloud-based collaboration services," in *2011 IEEE 5th International Conference on Internet Multimedia Systems Architecture and Application*, 2011.

[26] M. Ningning and C. Lizhen, "SMART: A platform for cross-domain business process collaboration," in *The 2010 14th International Conference on Computer Supported Cooperative Work in Design*, 2010.

[27] A. Salau, R. Dantu and K. Upadhyay, "Data Cooperatives for Neighborhood Watch," in *2021 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 2021.

[28] J. Krogstie and S. Gao, "A semiotic approach to investigate quality issues of open big data ecosystems," in *IFIP advances in information and communication technology*, 2015, p. 41–50.

[29] M. Zhang, F. Beltrán and J. Liu, "A Survey of Data Pricing for Data Marketplaces," *IEEE Transactions on Big Data,* vol. 9, no. 4, pp. 1038-1056, 2023.