Piano Protocol

A New Standard for Decentralized Data Licensing

Abstract

Piano Protocol is a decentralized data licensing infrastructure designed for the secure, transparent, and efficient exchange of high-value datasets. Initially focusing on healthcare data, Piano is extensible to any industry where data is sensitive, siloed, and crucial for innovation. As artificial intelligence and advanced analytics drive economic transformation, the availability of high-quality, ethically sourced data has become a primary bottleneck. Piano addresses the opacity, misalignment, and compliance challenges prevalent in today's data markets by providing programmable, auditable, and fair infrastructure for real-time data exchange. The protocol enables compliant licensing via smart contracts, featuring automatic royalty splits, immutable audit trails, and a governance model that aligns incentives across all participants. Piano's native token, $PIANO, is a soulbound governance token earned through active protocol participation, facilitating a progressively decentralized ecosystem. The go-to-market strategy prioritizes institutional onboarding with high-volume use cases like real-time Rx and claims data, establishing Piano as the foundation for the next generation of trusted, data-driven applications.

1. Introduction: The Vision for a Fair Data Economy

The global economy is increasingly powered by data, the critical fuel for artificial intelligence and analytical insights that drive innovation across sectors. From advancing medical research with real-world evidence to optimizing financial models and industrial processes, the demand for granular, reliable, and ethically sourced data is immense. However, the current data landscape often hinders progress, characterized by fragmentation, opacity, and misaligned incentives.

Piano Protocol envisions a global, transparent, and ethical data economy. We aim to

transform how data is licensed, valued, and utilized by establishing a new layer of digital infrastructure for programmable data exchange. Our vision is a future where data is treated as a yield-bearing asset class, where its value flows equitably to those who generate and curate it, and where programmable access replaces outdated brokerage bottlenecks. Piano is not merely a marketplace; it is the foundational infrastructure for trusted, data-driven applications, built on principles of transparency, utility-based pricing, and open design.

2. The Problem: Inefficiencies and Risks in Current Data Markets

Today's data markets are ill-equipped to meet the demands of the AI-driven economy, suffering from several critical flaws:

- **Opacity and Centralization:** A few dominant brokers control access to vast datasets, operating with limited transparency regarding data sourcing, usage rights, and pricing. This centralization creates bottlenecks and information asymmetry.
- **Misaligned Incentives:** Data originators (e.g., hospitals, research institutions, fintechs) often lack control over how their data is used and receive minimal compensation, if any, while intermediaries capture the majority of the economic value.
- **Barriers for Buyers:** Data buyers face inflexible pricing, delays in access, and uncertainty about data provenance and quality. This slows innovation and increases regulatory and operational risks.
- **Compliance and Privacy Risks:** Navigating complex regulatory landscapes (e.g., HIPAA, GDPR) is a constant challenge. The lack of auditable trails and clear consent mechanisms exposes both data sellers and buyers to significant compliance risks.
- **Stifled Innovation:** Emerging AI companies, researchers, and smaller developers are often priced out or excluded from accessing high-value datasets, hindering broader innovation.

These systemic issues mirror the inefficiencies of pre-cloud software: closed, costly,

and difficult to integrate. A new paradigm is needed to unlock the full potential of data assets in a secure, fair, and programmable manner.

3. Piano Protocol: The Solution for Decentralized Data Licensing

Piano Protocol offers a decentralized, permissioned infrastructure designed to facilitate the licensing and exchange of high-value datasets in a programmable, compliant, and fair manner. It empowers data owners to stream permissioned datasets and receive automated, usage-based compensation via smart contracts, while providing buyers with auditable, real-time access to data without opaque intermediaries.

Key features enabled by Piano Protocol include:

- Compliant, de-identified datasets licensed via smart contracts with automated royalty splits.
- Immutable audit trails tracing data access and enforcing usage limitations.
- Dynamic pricing capabilities that can adapt to data quality and buyer needs.
- Support for both high-throughput data streams (e.g., pharmacy transactions) and specialized datasets (e.g., rare disease registries).
- Optional SaaS tools for analytics, federated model deployment, and cohort discovery to enhance user experience.

Piano is architected to address the core bottlenecks in data licensing by focusing on institutional-grade supply, unbundling the dataset economy with flexible access models, and prioritizing compliance by design.

4. Core Principles of Piano Protocol

Piano Protocol is built upon the following core principles:

- **Transparency:** All licensing terms, data usage (in aggregate or de-identified forms), and revenue distributions are recorded and auditable on-chain, fostering trust among participants.
- **Fair Value Exchange:** Data originators and curators are fairly compensated based on the actual usage of their data, with automated royalty splits managed by smart contracts.
- **Compliance by Design:** The protocol integrates robust privacy and compliance mechanisms, supporting standards like HIPAA and GDPR, and enabling auditable data governance.
- **Data Sovereignty & Security:** Data owners retain control over their datasets, which remain in their secure, off-chain environments. Piano facilitates access and licensing without requiring data to be moved or centrally stored by the protocol.
- **Programmability & Composability:** Licensing agreements are encoded as smart contracts, allowing for flexible, automated enforcement of usage terms. The protocol's modular design enables third-party developers to build applications and services on top of Piano.
- **Progressive Decentralization:** Governance and control of the protocol will transition from Piano Labs (the initial developer) to the community of $PIANO token holders over time, ensuring long-term alignment with ecosystem needs.

5. Technical Architecture

Piano Protocol employs a hybrid architecture that combines the security and transparency of blockchain technology for governance and licensing with the flexibility of off-chain data storage and computation.

5.1. Hybrid Model: On-Chain Governance, Off-Chain Data

All raw datasets remain in secure, off-chain environments, either within the data provider's existing infrastructure or designated secure compute nodes operated by verified partners. This approach ensures data sovereignty, performance, and compliance with jurisdictional data residency requirements. On-chain components

manage the crucial aspects of discovery, licensing, access rights, and audit trails. Cryptographic hashes of dataset versions can be written to the blockchain to provide immutable proof of existence and versioning without exposing sensitive content.

5.2. In-Situ Data Management & Access Control

A core tenet of Piano is that data should, wherever possible, remain in-situ within the owner's trusted environment. The protocol facilitates access to this data through a sophisticated system of access control tokens (e.g., dataset NFTs or datatokens) and smart contracts. These tokens define the scope, usage rights, and duration of access for specific buyers or cohorts. This model allows data owners to license and monetize their assets without ceding physical control or duplicating large datasets, minimizing security risks and compliance burdens. Piano's architecture is designed to be infrastructure-agnostic, enabling existing data custodians to integrate seamlessly.

5.3. Smart Contracts: Licensing and Royalties

Smart contracts are the backbone of Piano's licensing and monetization capabilities.

- **Licensing Enforcement:** Each data access agreement is encoded as a smart contract defining strict usage controls, such as time-bound access, permitted use cases (e.g., model training only), data export limitations, and revocation rights. These terms are referenced and enforced programmatically.
- **Automated Royalties:** When data is licensed and payment occurs (in USDC or other approved stablecoins), the smart contract automatically distributes revenues according to a predefined split: typically 70% to the data provider, 29% to any facilitator or agent (e.g., node operator, marketplace application), and a 1-5% protocol fee to the Piano treasury. This ensures transparent and immediate compensation for all contributing parties.

5.4. Pricing Engine

Piano supports flexible pricing mechanisms to reflect data value and market dynamics:

- **Fixed-Rate Contracts (Phase 1):** Early deals may use fixed-price contracts defining flat-rate or per-record pricing for specified usage windows to ensure predictability.

- **Dynamic Pricing (Phase 2):** Over time, an AI-powered quality scoring engine can be integrated to drive dynamic pricing based on attributes like data freshness, density, uniqueness, and recency.

5.5. Privacy-Enhancing Technologies (PETs) and Compliance

Piano is designed with a compliance-first approach, incorporating features to handle sensitive data ethically and legally:

- **De-identification Standards:** Datasets must conform to established de-identification standards (e.g., HIPAA Safe Harbor or Expert Determination).
- **Privacy-Preserving Compute:** For highly sensitive data, Piano can support or integrate with federated learning, secure multi-party computation (SMPC), and compute-to-data environments, where algorithms are sent to the data rather than data being transferred.
- **Consent Management:** The protocol can enable on-chain tracking of consent and opt-out status, including versioned consent and tokenized consent receipts.
- **Regulatory Design:** Piano is built to meet requirements such as GDPR pseudonymization and access logging, GCP & 21 CFR Part 11 audit readiness for clinical datasets, and relevant financial data guidelines where applicable.

6. Tokenomics: The $PIANO Token

Piano introduces a native, non-transferable utility token—**$PIANO**—to govern protocol decisions and reinforce aligned participation.

6.1. Purpose and Design Philosophy

Unlike speculative assets, $PIANO is not meant for trading or direct financial gain. It is designed exclusively to reflect meaningful engagement with the protocol and to allocate governance power to active participants. Its core purpose is to facilitate decentralized decision-making regarding the protocol's evolution, parameters, and treasury.

6.2. Earning $PIANO

Each time a data licensing transaction occurs on the protocol, new $PIANO tokens are minted at a 1:1 ratio with the transaction's USD-equivalent value (settled in USDC or other approved stablecoins). The newly issued $PIANO tokens are split evenly between the participants of that transaction: 50% to the data buyer and 50% to the data seller. This structure directly ties governance influence to protocol usage and ensures that both supply-side and demand-side participants accrue a stake in the protocol's governance.

6.3. Token Properties

- **Soulbound:** $PIANO tokens are non-transferable and permanently bound to the wallet that earned them. This ensures that voting power cannot be bought or sold on secondary markets and must be earned through meaningful protocol engagement. This characteristic is also critical for enhancing the Sybil-resistance of the chosen voting mechanism.
- **Inflationary:** New $PIANO tokens are continuously created in direct proportion to economic activity on the protocol. This means the total supply of $PIANO will grow as protocol usage grows. This mechanism naturally dilutes the relative voting power of inactive holders over time, rewarding and prioritizing ongoing, active usage.
- **Decay for Inactivity (Burn Mechanism):** $PIANO tokens that remain unused in governance (i.e., have not been used for direct voting or have not been delegated to an active delegate) for a continuous period of 365 days are automatically burned. This policy prevents governance stagnation by removing dormant voting power from the system and ensures that quorum calculations reflect currently active or recently active participants.

7. Governance Model

Piano's governance model is designed to balance stakeholder influence with democratic safeguards, ensuring that voting power scales with protocol usage but cannot be easily or disproportionately dominated by any single entity.

7.1. Quadratic Voting (QV)

All on-chain governance decisions for Piano Protocol will be made using Quadratic Voting. In QV, the cost to cast multiple votes on a single proposal option increases quadratically. For example, casting 1 vote costs 1 unit of Vote Power, 2 votes on the same option cost 4 units (1+3), 3 votes cost 9 units (1+3+5), and so on. This mechanism allows participants to express the intensity of their preferences and enables large stakeholders to have proportional influence without easily monopolizing outcomes. It also ensures that strongly held minority opinions or broadly supported initiatives among smaller stakeholders can still achieve significant impact.

7.2. Delegation

$PIANO Vote Power can be delegated to another wallet address without transferring the soulbound $PIANO tokens themselves. This enables users, especially institutions or individuals who may not have the capacity to participate in every vote, to nominate trusted representatives to vote on their behalf. Delegation is revocable at any time by the token holder.

7.3. Participation Thresholds

- **Quorum:** A minimum of 25% of the currently active (non-burned) $PIANO tokens must participate in a vote for the result to be considered binding and valid. The burn mechanism for inactive tokens ensures that this quorum is calculated against a relevant base of engaged participants.
- **Passing Criteria:** A simple majority (>50%) of the total Vote Power cast (as weighted and calculated through the Quadratic Voting mechanism) is required for a proposal to pass. (Future consideration may be given to supermajority requirements for highly critical constitutional changes).

7.4. Rationale: Balancing Influence and Participation

This governance model intentionally allows larger buyers and sellers, who generate the most value and have the most at stake in the system, to accumulate more potential Vote Power. However, Quadratic Voting ensures that this accumulated power cannot be wielded monolithically, requiring broader consensus for significant changes and giving a stronger voice to widespread community preferences.

8. Progressive Decentralization and Protocol Stewardship

Piano Protocol is designed for a phased transition from initial centralized stewardship to robust community-led governance, ensuring stability in its early stages and true decentralization in maturity.

8.1. Phase 1: Foundational Stewardship by Piano Labs

At launch and during the initial development phase (Phase 1), Piano Labs (the founding company) will serve as the primary steward of the Piano Protocol. This centralized stewardship is crucial for agile development, rapid iteration, establishing initial partnerships, ensuring strong quality control for early data providers, and managing legal and compliance frameworks. During this phase, Piano Labs will guide key operational decisions and protocol upgrades.

8.2. Initial Governance Stake and Transition Plan for Piano Labs

To facilitate responsible stewardship and a gradual transition of control ("slow removal of training wheels"), the following mechanisms will be implemented:

- **Genesis Stake:** At the protocol's genesis, Piano Labs will receive an initial allocation of $PIANO tokens. This grant will be equivalent to 80% of the $PIANO tokens that would have been generated based on a reasonable projection of platform activity in its first year, or a similar mutually agreed-upon foundational grant. This provides Piano Labs with significant initial influence to ensure stability and adherence to the core vision.
- **Phased Influence Reduction:** For the first five years post-launch, Piano Labs will be eligible to receive programmatic, diminishing annual grants of new $PIANO tokens. The purpose of these grants is to allow Piano Labs to maintain an influential voice, targeting up to (but not necessarily exceeding) 50% of the total active voting power. The precise amount of these grants will be determined by a transparent formula based on the growth of overall network-generated $PIANO, ensuring the company's voting share reduces predictably over time as the

ecosystem's participation grows. This approach ensures Piano Labs can safeguard the protocol during its critical early years while transparently committing to a decentralized future.

8.3. Phase 2: Transition to Community-Led Governance

As the protocol matures, transaction volume increases, and the $PIANO token distribution becomes wider, governance decisions will increasingly shift to the community of $PIANO token holders through the on-chain Quadratic Voting mechanism. Data providers, buyers, developers, and other stakeholders will be empowered to propose, debate, and vote on protocol upgrades, fee structures, and other key parameters. Protocol smart contracts and administrative controls will progressively migrate to community-governed multi-signature wallets or fully on-chain, DAO-controlled upgrade paths.

8.4. Long-Term Vision: Data DAOs and Permissionless Ecosystem

In its most mature state (Phase 3), Piano envisions a vibrant ecosystem featuring specialized, delegated Data DAOs. These domain-specific groups (e.g., a rare disease registry DAO, a financial data DAO) could establish their own local policies, curate datasets, and manage specific data verticals within the broader Piano framework. The protocol will also move towards more permissionless onboarding of datasets and incentives for a wide range of ecosystem contributors, fostering a truly open and community-driven data economy. While Piano Labs transitions away from majority stewardship, it may continue to contribute as a significant participant and service provider within the ecosystem, potentially retaining a portion of protocol fees as a "creator fee" for its foundational work and ongoing contributions, subject to community governance.

9. Core Use Cases

Piano is designed for high-value, high-scrutiny data markets where transparency, auditability, and control are paramount. Initial use cases include:

- **Real-Time Rx & Claims Data:** Licensing de-identified pharmacy transactions and claims feeds to pharmaceutical teams, hedge funds, and market access groups. This provides immediate, recurring revenue and high-volume transaction flow. Data is de-identified under HIPAA Safe Harbor or Expert Determination frameworks.
- **Regulatory-Grade Real-World Evidence (RWE):** Enabling academic medical centers and research networks to license OMOP/CDM-compliant exports for pharmaceutical R&D, label expansions, and health economics research. Conforms to 21 CFR Part 11 for traceability.
- **AI Model Training on Multimodal Data:** Providing access to diverse datasets (ECG, imaging, genomics, sensor logs) for AI startups, research labs, and robotics teams, often utilizing federated learning or secure compute environments.
- **Synthetic Control Arms for Clinical Trials:** Repurposing legacy clinical trial data or observational cohorts to create cost-effective comparators for new studies.
- **Pharmacovigilance & Safety Monitoring:** Facilitating the submission and licensing of de-identified adverse drug event logs and patient-reported outcomes for safety monitoring by pharma and regulators.

10. Ecosystem and Monetization

Piano generates sustainable revenue while enabling efficient transactions for data providers and buyers.

10.1. Protocol Fees

A small transaction fee (1-5%) is applied to each data license executed via the protocol. This fee is collected in USDC (or other approved stablecoins) at the point of transaction. Typically, this fee is allocated to the Piano Treasury (controlled by $PIANO token holders in later phases) to fund ongoing protocol development, governance operations, and ecosystem grants.

10.2. Optional SaaS Layers

Piano Labs (and potentially third-party developers in the future) may offer optional

SaaS tools that enhance the protocol's utility for enterprise users. These could include:

- Analytics and reporting dashboards.
- Pricing APIs and market intelligence tools.
- Federated model management infrastructure.
- Custom cohort builders. These services would follow a separate, usage-based pricing model.

10.3. Role of Data Providers, Buyers, and Facilitators

- **Data Providers (Sellers):** Hospitals, research organizations, pharmacies, fintechs, etc., can license their data, set terms, and earn automated royalties (typically 70% of the transaction value).
- **Data Consumers (Buyers):** Pharma companies, AI startups, hedge funds, researchers, etc., gain programmatic access to high-quality, auditable data.
- **Facilitators/Agents:** Entities that operate nodes, curate data, build applications on Piano, or otherwise facilitate transactions can earn a share of the transaction value (typically 29%).

11. Roadmap and Go-To-Market Strategy

Piano's go-to-market plan is deliberately staged to build liquidity, credibility, and utility without relying on speculation.

- **Phase 1: Anchor Supply with Real-Time Rx and Claims Data.** Launch by onboarding high-throughput, high-demand de-identified datasets with clear legal frameworks (HIPAA Safe Harbor) and existing buyer demand (pharma, hedge funds). This provides immediate utility and revenue.
- **Phase 2: Expand to Research-Grade Datasets via Institutional Nodes.** Onboard academic or research hospital systems to contribute curated RWE bundles (e.g., OMOP/CDM exports), demonstrating viability for higher-margin, regulatory-grade data. This phase also tests more complex governance and compliance requirements.
- **Phase 3: Launch Open Listings & Developer SDKs.** Introduce an open listing portal for institutional data providers and release developer SDKs to encourage

third-party application development and expansion into new verticals beyond healthcare. This marks the shift towards a more decentralized ecosystem.

12. Risk Factors and Mitigations

Piano Protocol acknowledges and addresses potential risks through its architecture and governance:

- **Sybil Attacks & Fake Data:** Initial permissioned onboarding of verified institutional providers. Future validation via metadata heuristics, cross-referencing, and potential AI-driven anomaly detection. Soulbound $PIANO tokens also mitigate Sybil attacks in governance.
- **Data Misuse by Buyers:** Smart contracts encode strict licensing terms (scope, duration, revocability). Immutable access logs support audits. Compute-to-data options limit raw data exposure for high-risk scenarios.
- **Regulatory Uncertainty:** Compliance-first design, adherence to existing standards (HIPAA, GDPR), and support for PETs ensure adaptability to evolving regulations. Data remains off-chain and in local jurisdictions where appropriate.
- **Cold Start Dynamics:** Strategy focuses on specific, high-value datasets with known demand (Rx/claims) to create immediate liquidity and utility, rather than abstract monetization pitches.
- **Ethics & Public Perception:** Emphasis on data owner control, transparent compensation, auditable consent (where applicable), and clear terms of use. Focus on de-identified data for many core use cases.

13. Conclusion

Piano Protocol is more than a data marketplace; it is foundational infrastructure for a new data economy—one built on trust, transparency, fair value exchange, and programmability. By addressing the critical limitations of current data markets and providing institutions with a compliant, efficient, and ethical way to license and utilize high-value data, Piano unlocks new opportunities for innovation in AI, healthcare, finance, and beyond.

The carefully designed tokenomics of the $PIANO governance token, combined with a

robust Quadratic Voting mechanism and a clear path for progressive decentralization, ensures that the protocol will evolve in alignment with the interests of its active participants. Piano Protocol offers the rails for the next generation of data-driven applications, fostering an ecosystem where data is treated as a dynamic, yield-bearing asset, accessible and priced by utility, and open by design.