Quelldata: Decentralized Data Crowdsourcing Platform for AI

Whitepaper v1.2 – May 2025

Legal Disclaimer Nothing in this White Paper constitutes an offer to sell, or the solicitation of an offer to buy, any tokens. Quelldata is publishing this document solely for informational purposes and to solicit feedback from the public regarding its proposed platform architecture, token design, and governance model. If and when Quelldata offers any tokens including but not limited to the Quelldata Token (QLD) or any derivative instruments such as SAFTs, it will do so through formal offering materials, including risk disclosures, legal frameworks, and compliance documentation. These materials will supersede this document and may differ materially from the information presented herein. Any token offering conducted in the United States or other regulated jurisdictions will be made only to qualified or accredited investors, in full accordance with applicable securities laws. This document does not constitute investment advice, legal guidance, or a guarantee of project success. Quelldata's forward-looking statements are based on current plans and assumptions, which are inherently subject to risks, uncertainties, and external variables including but not limited to technological developments, market adoption, and regulatory changes. No representation is made regarding the future utility or value of the QLD token.

Abstract

Abstract This whitepaper introduces Quelldata — a decentralized, tokenized platform designed to crowdsource and validate ethically sourced datasets for AI development. By combining Web3 infrastructure, smart contracts, and peer-to-peer data verification protocols, Quelldata seeks to democratize access to high-quality training data. Central to the platform is the Quelldata Token (QLD), a non-inflationary utility token deployed on the Solana blockchain, enabling direct incentives for data contributors and validators. The platform utilizes a "Proof-of-Quality" mechanism enforced by a decentralized autonomous organization (DAO), ensuring data integrity through randomized peer-review, zero-knowledge proof validation, and bias detection modules. Quelldata's architecture leverages a hybrid stack: on-chain governance and payments via Solana's high-speed network, decentralized storage through IPFS and Arweave, and an AI engine for continuous data quality analysis. The result is a trust-minimized

pipeline that enables researchers, startups, and enterprises to request, source, and verify datasets with high granularity and transparency. We explore a range of real-world use cases spanning healthcare, agriculture, and ethical AI and present a tokenomics model designed for long-term ecosystem sustainability. Through community participation, open infrastructure, and transparent incentives, Quelldata redefines how data is ordered, validated, and earned in the age of decentralized intelligence.

1. Introduction

Quelldata is a decentralized infrastructure for sourcing, validating, and distributing high-quality datasets for artificial intelligence systems. In response to the growing demand for transparent, ethical, and manipulation-resistant training data, Quelldata combines blockchain technology, decentralized storage, staking mechanisms, and token-based incentives to establish a new standard in data accessibility.

Conventional data acquisition models are centralized, opaque, and lack equitable mechanisms to compensate contributors and validators. Quelldata redefines this process through an open, verifiable, and scalable network in which each participant contributes to a fairer and more resilient data economy for AI.

2. Problem Statement

Existing data supply chains suffer from structural inefficiencies. Centralized providers, such as Scale AI or Appen, monopolize access to data. The labor behind dataset curation is underpaid and unrecognized, often executed under exploitative conditions. There is no systemic transparency regarding how data is collected, nor assurance that it adheres to regulatory or ethical standards.

Current frameworks also expose AI developers and enterprises to regulatory risk, particularly with the enforcement of GDPR, HIPAA, and related data protection laws. Furthermore, the dependency on proprietary APIs constrains innovation and limits the development of localized AI applications in emerging markets.

3. Solution Overview

Quelldata introduces a permissionless, modular architecture where data contributors upload labeled data across modalities such as images, text, and audio. Validators with on-chain reputation scores confirm or reject submissions through peer-reviewed staking rounds. Data clients, including researchers, startups, and institutions, create bounties via smart contracts funded in QLD tokens.

A multi-layered reputation system and a proof-of-quality protocol ensure data integrity while aligning incentives. A decentralized autonomous organization (DAO) governs parameter adjustments, fund allocations, and protocol evolution.

Quelldata positions itself as an open-weight infrastructure layer, composable with AI pipelines, data marketplaces, and LLM training ecosystems..

4. System Architecture

4.1 Core Components

Frontend:
Quelldata provides a clean Web2 interface designed for intuitive task navigation. Users can submit data (e.g. images, videos), label attributes, verify other submissions, manage staking, and monitor their reputation — all without requiring Web3 experience.

Smart Contracts:
Deployed on Solana, the smart contract layer governs the core protocol logic: request creation, staking mechanics, reward distribution, validator coordination, and DAO governance. Tasks are uniquely represented as on-chain challenges.

Storage Layer:
Raw user-submitted data is temporarily stored in decentralized storage (e.g., IPFS). Finalized and verified datasets are archived for long-term access and licensing. The system ensures anonymity and compliance (e.g., GDPR-ready exports for B2B use).

Validation System (AI + Human):
Data quality is ensured through a hybrid model:

- AI agents automatically detect duplicates, noise, and format inconsistencies
- Human validators are required to reach consensus on task results
  This system filters out low-quality or malicious contributions, ensuring real-world usability

Reputation Layer:
All actions — submissions, validations, disputes — are recorded on-chain with cryptographic proof. Each contributor builds a non-transferable reputation profile in the form of an NFT. Higher reputation increases access to high-value tasks and boosts rewards.

4.2 Proof-of-Quality Mechanism

Each task requires a minimum quorum of human validators to reach consensus.
Participants must stake QLD tokens as collateral when submitting or verifying data. If a

participant behaves dishonestly or provides low-quality output (as determined by consensus and AI agents), their stake is partially or fully slashed, and their reputation score is penalized. This incentivizes high-quality contributions while eliminating the need for centralized moderation.

4.3 Incentivization Layer

Quelldata uses a gamified engagement loop to drive long-term participation:

- Stake-based rewards scale with contribution volume, accuracy, and task priority
- Reputation multipliers grant top contributors early access to premium challenges
- Referral and streak bonuses further improve retention over time

This ensures a self-sustaining ecosystem, where data quality improves as the network grows.

5. Governance

Governance of the Quelldata protocol is managed by the community through a DAO. Holders of QLD tokens may create and vote on proposals related to system changes, funding allocation, and policy updates. The DAO employs quadratic voting to ensure fair influence distribution, and all votes are executed via smart contracts with timelocks and multisig fallback.

Key governance domains include:

- Approval of new task types and categories
- Allocation of treasury funds and grants
- Rules for reputation scoring and penalties
- Partner onboarding and DAO-driven incentives

6. Roadmap

Phase	Date	Milestone
1. MVP	Q2 2025	Launch of web platform with crowdsourced data tasks & IPFS integration
2. Mobile App	Q3 2025	Android/iOS app + data marketplace activation
3. AI Agents	Q4 2025	Deployment of autonomous agents for task generation and pre-validation
4. LLM Plugin	Q1 2026	Launch of Quelldata plugin/API for Large Language Models
5. Private B2B Suite	Q2 2026	Deployment of GDPR-compliant Quelldata nodes for enterprises
6. Monetization Tools	Q3 2026	Anonymous dataset monetization layer for both users and organizations

7. Compliance & Risk Mitigation

Quelldata is designed with compliance and risk mitigation at its core. The platform natively supports GDPR and CCPA requirements, offering tools for data anonymization and deletion upon request. Sensitive data is processed using zero-knowledge proofs to ensure privacy without compromising validation. The project operates under a legally registered foundation in Switzerland, structured in accordance with digital ledger technology (DLT) regulations. All smart contracts undergo rigorous third-party security audits to ensure protocol integrity. Additionally, identity verification procedures such as KYC and AML are enforced for all grant recipients and high-value withdrawals to maintain regulatory alignment and prevent misuse.

8. Conclusion

Quelldata is not merely a platform but a movement to reclaim the data economy for the global community. Through decentralized infrastructure, open participation, and verifiable quality, we empower the next generation of AI development. Join the future of fair data. Contributors get rewarded for quality data, clients source unique datasets at lower costs, and builders integrate with our open API and DAO. The future of AI is shared. Let’s build it—together.

Contact & Links quelldata.com Twitter Discord Telegram Disclaimer: QLD is a utility token. This document is not a prospectus or solicitation for investment.

9. References

1. Benet, J. (2014).
   IPFS - Content Addressed, Versioned, P2P File System. arXiv preprint arXiv:1407.3561.
   (Decentralized storage architecture)
2. Zyskind, G., Nathan, O., & Pentland, A. (2015).
   Decentralizing Privacy: Using Blockchain to Protect Personal Data. IEEE Security and Privacy Workshops (SPW), 180–184.
   (Privacy and compliance in blockchain)
3. Koutroumpis, P., Leiponen, A., & Thomas, L. D. W. (2020).
   Markets for Data. Industrial and Corporate Change, 29(3), 645–660.
   (Data markets and incentive models)
4. Hashemi, S. A., Faghri, F., & Faghri, A. (2021).
   Decentralized Data Marketplaces: A Review. IEEE Access, 9, 56194–56211.
   https://doi.org/10.1109/ACCESS.2021.3071856
   (Decentralized data exchange platforms)
5. Wang, H., Xu, Z., Fujita, H., & Liu, S. (2020).
   Towards trustworthy AI: Blockchain-based architecture and case study. Information Fusion, 58, 341–350.
   (Blockchain for AI data integrity and trust)
6. Huckle, S., & White, M. (2016).
   FinTech and the technological singularity: A blockchain-based decentralized autonomous organization. International Journal of Blockchains and Cryptocurrencies, 1(1), 1–12.
   (Decentralized governance/DAO)