AI & SOCIETY
https://doi.org/10.1007/s00146-024-02166-w

OPEN FORUM

Anomaly detection and facilitation AI to empower decentralized autonomous organizations for secure crypto-asset transactions

Yuichi Ikeda[1] · Rafik Hadfi[2] · Takayuki Ito[2] · Akihiro Fujihara[3]

Received: 14 May 2024 / Accepted: 16 December 2024

Abstract
This proposal introduces a novel decision-making framework to advance safe economic activities in cyberspace. We focus on identifying anomalies within crypto-asset trading, recognized as potential sources of criminal activity, severely undermining the credibility of such assets. Detecting and mitigating such anomalies holds significant societal implications, particularly in fostering trust within blockchain networks. We aim to bolster the "social trust" inherent to blockchain technology by facilitating informed economic activities in cyberspace. To achieve this, we propose integrating two artificial intelligence (AI) systems into a blockchain-based decentralized autonomous organization (DAO). The first AI application involves amalgamating various anomaly indicators, spanning from cluster coefficient, entropy, triangular motif analysis, correlation tensor analysis, loop component by Hodge decomposition, loop causality detection, network classification using graph Laplacian, and persistent homology analysis, into a comprehensive indicator using a Boltzmann machine. The second AI application entails deploying conversational AI to guide and support traders, aiding them in making informed trading decisions. This system is designed to alert DAO members to anomalies based on the integrated indicators, especially during massive price fluctuations. We operate under the assumption of close collaboration between governments, experts, traders, system developers, and operators to effectively organize DAOs. The primary technical challenge in our proposal lies in developing a wallet assisted by an intelligent software agent capable of safe interactions with traders within a unified DAO. With this organization, we envision fostering a global economic ecosystem where physical and cyber worlds converge, allowing democratic economic participation.

Keywords Markets · Trading · Anomaly detection · Blockchain · Smart contracts · Decentralized autonomous organization · Cyber-physical systems · Token economy · AI · Conversational AI · Agents · Multiagent systems · Natural language processing · Boltzmann machine · Graph Laplacian · Correlation tensors · Democracy · Governance

1 Introduction

In today's era of digital and economic globalization, capitalist economies are exploring new frontiers. One of the emerging frontiers lies in integrating digital economic activities into the regulated framework of the international economy, creating what could be perceived as a cyber-physical integrated economy. For instance, economic activities in cyberspace are now facilitated by blockchain technology, serving as the foundation for crypto-assets (Yano et al. 2020). Within this digital domain, novel organizational structures are emerging, distinct from traditional enterprises characterized by hierarchical management and centralized control over goods and services distribution. The rise of decentralized autonomous organizations (DAOs) constitutes a shift from conventional organization models in replacing or supplementing traditional ones (Altaleb 2022). DAOs offer a decentralized approach for collaboration and decision-making beyond hierarchical top-down directives. This transformative potential holds promise for cultivating societies cantered on wellbeing and sustainability and departing from the narrow emphasis on GDP-driven economic growth.

✉ Yuichi Ikeda
ikeda.yuichi.2w@kyoto-u.ac.jp

[1] Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University, Kyoto, Japan

[2] Graduate School of Informatics, Kyoto University, Kyoto, Japan

[3] Graduate School of Engineering, Chiba Institute of Technology, Chiba, Japan

Published online: 07 January 2025

The spread of social media has resulted in an unprecedented volume of information that overwhelms the cognitive capacities of individuals. It has also precipitated a shift in societal dynamics from democracy towards populism in many countries (Berman and Snegovaya 2019). Democracy, a cornerstone of capitalist economies, holds a well-documented significance for all sorts of economic activities. We are unfortunately witnessing an increase in anomalies within the context of cyber-physical economic activities such as crypto-asset trading. These transgressions undermine the credibility of crypto-assets, posing significant threats to their adoption by the larger public. Consequently, the detection and prevention of such crimes carry substantial social importance. Recent years have witnessed a surge in the exploration of mathematical methodologies for identifying anomalies in crypto-asset transactions (Ikeda 2022; Ikeda and Chakraborty 2023; Chakraborty et al. 2023, 2024; Aoyama et al. 2022), with observable instances of practical application within corporate settings. Moreover, there have been propositions for leveraging artificial intelligence (AI) within online social platforms to foster democratic decision-making that uphold individual autonomy (Hadfi 2022; Ito et al. 2022; Hadfi et al. 2023;).

The challenges raised so far require the involvement of various actors. Bridging the gap between academic research and its practical application within society, involving a spectrum of stakeholders including citizens, governments, and universities, poses numerous challenges. A novel approach for leveraging information technology to enact the findings of academic research within society is depicted in Fig. 1. In Fig. 1a, the process begins with AI-driven predictions based on mathematical models developed within university settings by scientists, forecasting the occurrence of specific events. Subsequently, Fig. 1b illustrates a democratic consensus-building process facilitated by AI, wherein stakeholders engage in discussions to formulate measures addressing the predicted events, using a wallet assisted by an intelligent software agent. Finally, in Fig. 1c, the control phase unfolds as smart contracts are automatically executed in response to the decision of the traders and the anomalies. A smart contract is generally defined as a self-executing digital agreement signed and stored on a blockchain network, and activated automatically when its terms and conditions are fulfilled. This deployment is also facilitated by an oracle, serving as an external input mechanism, thereby ensuring effective control measures are swiftly enacted in response to the anticipated events.

Vitalik Buterin developed a nice framework for the different types of AI integration with blockchain and the relation to crypto-economics, summarized in the following four categories: (1) "AI as a player in a game", (2) "AI as an interface to the game", (3) "AI as the rules of the game", and (4) "AI as the objective of the game" (Buterin 2024). The intelligent wallets proposed in Fig. 1b fall into category (2) "AI as an interface to the game", while the anomaly detection AI proposed in Fig. 1a and the facilitation AI proposed in Fig. 1c are category (3) "AI as the rules of the game". In addition, AI in category (4) "AI as the objective of the game" brings up the topic of how the two AIs in our proposal are trained. The AI integration with blockchain should ideally be transparent and open. However, as Buterin pointed out, this might make them vulnerable to adversarial attacks.

As depicted in Fig. 1, our proposal advocates for (a) predictive, (b) consensual, and (c) controlled economic activities within cyberspace. To this end, we propose to integrate two distinct AI systems. The first AI synthesizes anomaly indicators using various mathematical methodologies. The second AI offers guidance and support to traders in making informed decisions amidst anomalous trading conditions. Integrating these two functionalities, the system alerts the DAO members to anomalies synthesized from indicators, particularly during price spikes. Our overarching objective from this proposal is to promote public trust in blockchain technology and catalyze the advancement of a cyber-physical integrated economy. We choose to build such trust through the use of rigorous mathematical methodologies developed by scientists within an academic framework.

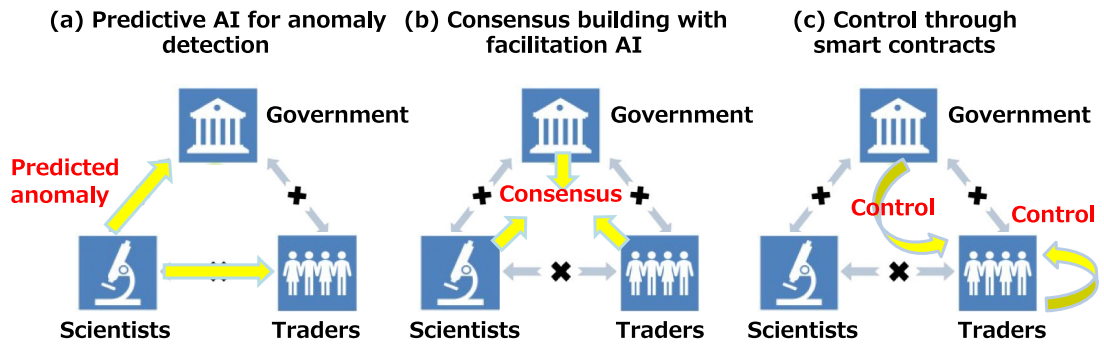

Fig. 1 Mechanisms for translating academic research findings into societal applications

Next, we start by describing our vision of DAOs since they constitute the building block of our envisioned society. In Sects. 3 and 4, we identify the type of anomalies that threaten DAOs and provide mathematical ways to detect such anomalies. In Sect. 5, we illustrate the use of AI to facilitate the decisions of the DAO members. In Sect. 6, we characterize the governance aspects of the DAO through smart contracts. In Sect. 7, we conclude by summarizing our findings.

2 Blockchain-based decentralized autonomous organization

Web 3.0 is revolutionizing the dynamics of information by leveraging blockchain technology to redefine its creation, ownership, and distribution mechanisms. This transformative paradigm shift is additionally creating novel values that affect all aspects of society. First, the notion of ownership within Web 3.0 introduces ground-breaking concepts. nonfungible tokens (NFTs), akin to digital deeds, emerged in 2022, catalyzing a remarkable $41 billion market. On the other hand, DAOs are operating in a manner akin to corporations but are devoid of a central authority (Altaleb 2022). They can democratically allocate funds and make decisions via member voting, enforced by cryptographic protocols. Such mechanisms are prone to occasional issues, such as the whale problem or collusion in DAO voting mechanisms. These issues can be mitigated by implementing robust voting mechanisms (Tamai and Kasahara 2024). Voting mechanisms can even be automatically evaluated using multiagent systems and machine learning techniques (Hadfi and Ito2023). Moreover, decentralized finance (DeFi) is experiencing a rise in popularity, championing a financial ecosystem independent from traditional banking institutions. This growing movement advocates for funding mechanisms liberated from the constraints of conventional banking systems.

Consider the software industry as an illustrative example, comparing the operational dynamics of a traditional stock company with those of a DAO. In a stock company, governance relies on resolutions made during shareholder meetings, whereas a DAO operates through community-driven agendas, fostering discussion in a threaded format. Following community consensus, proposals are drafted on behalf of the DAO, and token holders subsequently vote on these proposals. While the assets of a stock company primarily revolve around its proprietary products, a DAO functions as a public protocol, generating socially significant public goods within cyberspace. Unlike stock companies, which typically keep their source code confidential, DAOs uphold transparency by publicly disclosing their source code as open-source software. Moreover, while stock companies often aim at market dominance and profits, DAOs prioritize collective creation and ecosystem development, distributing profits among multiple contributors rather than concentrating them within a single entity. In terms of data management, stock companies tend to store information within internal databases, limiting accessibility. In contrast, DAOs leverage blockchain technology to prioritize openness and verifiability as fundamental tenets of their operations. Finally, participation in stock companies tends to be restricted and exclusive, primarily limited to employees or shareholders. Conversely, DAOs embrace inclusivity, welcoming contributions from anyone interested in participating and contributing to the community's endeavours.

Examples of DAOs encompass diverse social structures or organizations, spanning from investment-oriented endeavours such as fundraising, crowdsourcing, and charitable initiatives to ventures involving crypto tokens or NFTs trading and investment (Wattenhofer et al. 2024, Altaleb 2022). Additionally, many DAOs are actively engaged in software engineering, contributing to the development, modification, and maintenance of open-source software infrastructure, including blockchain systems and decentralized financial applications. However, the definition and structural possibilities of DAOs remain ambiguous, prompting inquiries into their characteristics and organizational forms. The Legal Committee's call for evidence seeks to gather perspectives on this emerging organizational paradigm. Presently, while numerous DAOs exist, only a handful are structured within the confines of national laws. This poses pertinent questions regarding the legal standing of DAOs, the obligations of their participants, and the regulatory frameworks governing their operations. In the dynamic landscape of DAOs, a significant amount of value is generated, exchanged, and occasionally lost. These activities underscore the pressing need to clarify the legal status of DAOs and delineate the regulatory landscape applicable to them.

DAOs offer several distinct advantages. First, they excel in efficient funding mechanisms, leveraging smart contracts to streamline fundraising processes, thus outperforming traditional organizational structures. This capability enables DAOs to secure funding expediently from a global pool of contributors, representing a substantial advantage. Second, DAOs provide clear and meaningful incentives for work. Community members engaged in a DAO can contribute to alignment with their preferences and expectations for compensation, fostering a sense of ownership and commitment. This feature holds promise as a potent tool for fostering internal ventures within the organization. Third, DAOs exhibit swift startup and operational agility, allowing rapid deployment and action. This agility facilitates adaptability in response to evolving circumstances, positioning DAOs as dynamic entities primed for innovation and responsiveness in the ever-changing landscape of cyberspace economics.

In this proposal, we advocate DAOs for their ability to organize digital communities intelligently and fulfil a mission of significant social importance. Our conception leverages various governance tokens to fine-tune incentives for participants, fostering a system where individuals are duly rewarded for their contributions to the community. Decision-making authority is decentralized, ensuring all DAO members possess equal influence (Altaleb 2022). The decisions pertaining to fund utilization and management within the organization are enacted through smart contracts upon attaining a majority consensus among voting members. DAOs exemplify a novel breed of online, multi-participant organizational structures, underpinned by software systems like blockchain and smart contracts. With their rising prominence within the context of crypto-assets and decentralized finance, DAOs serve as pivotal components of these ecosystems. It is important to note that the term "DAO" encapsulates a spectrum of organizational structures, with their legal classification contingent upon specific structural arrangements. Many DAOs leverage smart contracts to automate facets of their internal operations, embodying a dynamic evolution in organizational governance.

The primary drawback associated with DAOs is their protracted coordination time for decision-making. Nonetheless, they offer the distinct advantage of decentralization, with most DAOs fostering democratic operations through platforms such as Discord, promoting an egalitarian exchange of viewpoints. However, this democratic decision-making process often proves inefficient for organizational efficacy, necessitating facilitation—where AI proves to be a particularly effective tool (see Sect. 4). A secondary challenge stems from the jurisdictional complexities surrounding blockchain-based systems, as they often operate beyond the purview of existing legal frameworks. Legislation pertaining to DAOs remains underdeveloped in many countries and regions, lagging the rapid evolution of these decentralized structures. Unlike traditional organizational models, DAOs function based on democratic principles and systemic operational procedures, highlighting the imperative for legal frameworks that accommodate their unique structure. Collaboration between users, system developers, and operators underscores the need for legal support tailored to the DAO organizational paradigm. Furthermore, careful consideration must be given to the nature of user participation, as the mere receipt of warning services may not suffice to confer membership within a DAO. The clarification of participation criteria is essential to ensure the effective functioning and legitimacy of these decentralized entities.

A simple example will be used to illustrate how arguments among members can arise in DAO decision-making. In this paper, we consider a DAO consisting of traders who wish to trade crypto assets securely, scientists who believe they can predict the occurrence of crypto asset price spikes and crashes, and government regulator officers who wish to oversee crypto asset trading. To become a member of this DAO, it is necessary to purchase a governance token when joining and we assume a typical DAO in which the opinions of members with more tokens have more influence in the DAO's decision-making. The novelty of this study is that it uses not only SNS such as Discord but also consensus facilitation AI for discussions among members. Let us consider the following two situations. First, suppose that scientists detect that a bubble is about to occur using the anomaly detection AI (Sect. 4), and that scientists recommend that traders refrain from trading crypto assets for the time being to avoid traders' losses from a price collapse, since traders are not involved in the price spike. The scientists recommend that traders refrain from trading crypto assets for the time being to avoid trader losses due to price crashes. As a result, traders will be divided into two groups: those who think they should continue trading and those who think they should refrain from trading. A discussion will begin between the two groups, and the facilitation AI (Sect. 5) will bring the traders to an agreement. Next, consider a situation in which it is not sufficient for traders to refrain from trading crypto assets to avoid losses due to a price collapse: the DAO decides to adopt smart contract control using financial options (Sect. 6) to prepare for the occurrence of a bubble in advance. There would be a group of traders who believe that risk aversion using financial options is effective, a group that believes that financial options are unnecessary, and a group of government regulator officers who would prefer those transactions to be held off for the time being. Different points of view will be discussed among these members, and a facilitation AI will lead them to agreement.

3 Anomaly

An anomaly, or anomalous transaction, refers to financial asset transactions exhibiting characteristics significantly divergent from what is deemed "normal", indicative of aberrant trading behavior. The determination of anomalous status depends on contextual factors and fluctuates with specific transactional and market conditions. Financial transactions encompass various characteristics, with anomalies perceived differently depending on the metrics employed. A transaction flagged as anomalous under one set of criteria may not meet the same designation under an alternative framework. Regulatory oversight, embodied by entities such as Japan's Financial Services Agency (FSA), is tasked with scrutinizing these anomalies within financial transactions and implementing requisite interventions (Takefuji 2023; Reurink 2019).

Anomalies within crypto-asset transactions, often indicative of criminal activities, pose a significant threat to the

credibility of such assets. Hence, the social importance of detecting these anomalies cannot be overstated. Currently, companies involved in crypto-asset trading, particularly in detecting crimes like money laundering, rely on manual and individualized responses to identify anomalies. However, there exists a pressing societal need for automating anomaly detection processes. Financial institutions and exchange market operators routinely report various anomaly events to regulatory bodies such as the FSA in response to regulatory mandates. It is conceivable that the volume of reported anomalies in Japan surpasses that of other countries. Nonetheless, the effectiveness of utilizing these reported anomalies by the FSA may be hindered by the varying quality of the reports submitted.

Anomalies within financial transactions often signal potential illicit activities such as money laundering, arbitrage, and fraud (Valla et al. 2023; Hilal et al. 2022; Thudumu 2020; Ahmed et al. 2016). These anomalies can manifest in various forms, including unusually high trading volumes preceding major announcements, suggestive of insider trading. Furthermore, anomalies may coincide with abrupt price spikes or crashes, indicating market manipulation or other irregularities. While some anomalies may have minimal impact on prices, significant fluctuations typically hint at underlying irregularities. Thus, scrutinizing transaction data during periods of pronounced price movements becomes crucial. Below, we illustrate seven common types of anomalies:

1. *Front-running*: occurs when traders exploit advance knowledge of others' orders for personal gain, such as using client order information to execute profitable trades.
2. *Pump and dump*: involves artificially inflating the price of a low-value asset through substantial investments, followed by a coordinated sell-off to profit from misled investors.
3. *Wash trading*: a manipulation tactic where traders execute fake trades with themselves to fabricate market activity, often without any actual transfer of funds, to create a false impression of market health or to generate broker commissions.
4. *Cross trading*: involves trading the same contract month across different accounts of the same owner, potentially leading to market manipulation and legal repercussions.
5. *Insider trading*: occurs when individuals trade securities based on non-public, material information, typically resulting in unfair gains for those with privileged knowledge.
6. *Airdrop*: a marketing strategy where crypto-assets or tokens are distributed free to users, sometimes as part of a promotional campaign. It is a useful tool to market a new token. However, airdrops can be misused to drain the wallets of naive users.

However, identifying the specific transactions corresponding to these anomalies remains challenging due to the limited trader information available in transaction data, especially within the context of crypto-assets. Despite detailed analysis, distinguishing individual anomalies from transaction data proves elusive, primarily because trader identities are obscured, with only cryptographic hash values of source and destination wallets provided. In our proposal, we concentrate on periods characterized by significant price fluctuations. Through the analysis of transaction data utilizing multiple anomaly indicators, we aim to classify transactions that deviate from normal transactions as anomalies. Another lurking challenge that needs to be addressed, is the possibility that the anomalies (1–6) are caused or accentuated algorithmically using trading bots. Such bots utilize algorithms to automatically execute buy and sell orders on exchanges, often resulting in heightened price volatility due to similar trading patterns among multiple bots.

4 Anomaly detection AI

Anomaly detection AI, or alternatively predictive AI, is defined as a sequence that unfolds in three steps as outlined in Fig. 2. The dataset under analysis encompasses not only authentic transactional records of crypto-assets spanning both bubble and normal periods but also synthetic data. This synthetic dataset incorporates anomalous events representative of the seven types delineated in the preceding section, seamlessly interwoven with authentic transactional data from normal periods of activity.

(1) *Step 1*: network time series derived from crypto-asset transaction data undergo multifaceted analysis to derive network feature time series, capturing transaction anomalies. Concurrently, fluctuations in crypto-asset prices across exchanges are scrutinized to generate fluctuation feature time series, indicative of price anomalies. Through a comparative assessment between the transaction anomalies in network feature time series and the price anomalies in fluctuation feature time series, it is possible to identify the anomaly indicators that are more likely to influence price fluctuations. The left segment of Fig. 2 illustrates Step 1. Mathematical methodologies to estimate various features for transaction and price fluctuation anomalies and their empirical evaluation will be reported elsewhere (Ikeda et al. 2024).

(1–1) *Network feature time series depicting transaction anomaly*: employing graph theory, we scrutinize

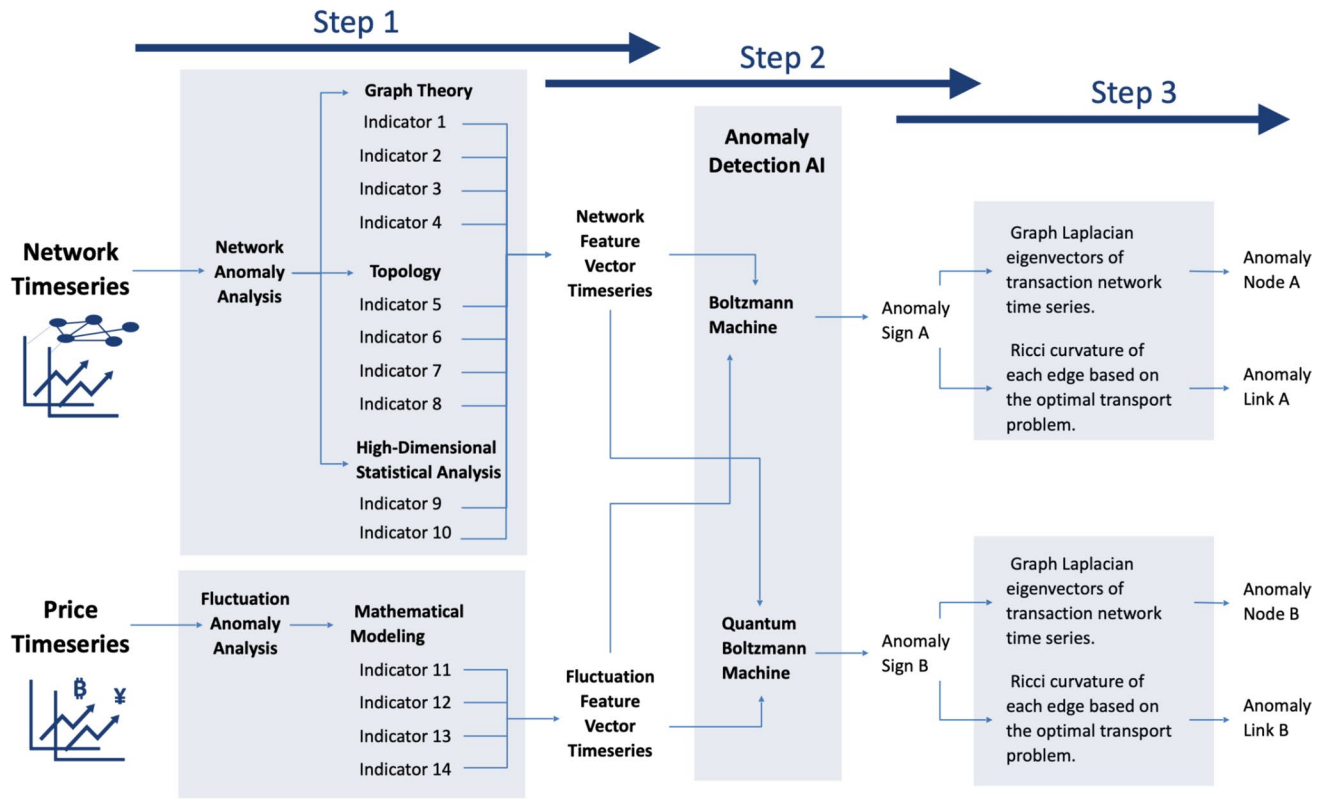

Fig. 2 Sequence of anomaly detection AI

various indicators: [Indicator 1] cluster coefficients, [Indicator 2] entropy, [Indicator 3] triangular motifs (comprising different motif types), and [Indicator 4] transaction loops, considering the temporal occurrence of edges. Through a topological lens, we delve into [Indicator 5] transaction loop components via Hodge decomposition, [Indicator 6] graph Laplacian eigenvalue distance classification (incorporating eigenvalues and eigenvectors), [Indicator 7] topological data analysis (including the number of transaction loops and Betti numbers), and [Indicator 8] Ricci curvature based on optimal transport problems (encompassing mean curvature and edge-specific curvatures). Additionally, we conduct high-dimensional statistical analyses, encompassing [Indicator 9] correlation tensor analysis and [Indicator 10] feature extraction of transaction frequency statistics (including the number of nodes within each branch).

(1–2) *Time series analysis of price anomalies*: in our examination of inter-exchange price fluctuations, we propose to analyze time series data derived from minute-by-minute trading activity. Specifically, we compute the ratio of high to low prices within a moving time window. This process allows us to extract the maximum fluctuations characteristic of four distinct time scales: minute, hour, day, and week. Subsequently, mathematical modelling is conducted within the following time windows: [Indicator 11] for minutes, [Indicator 12] for hours, [Indicator 13] for days, and [Indicator 14] for weeks.

(2) *Step 2*: this pivotal step aims at utilizing feature vectors comprising diverse anomaly indicators to provide a comprehensive assessment that determines whether an event qualifies as an anomaly. Specifically, individual indicators may vary in their anomaly classification. An AI is, therefore, tasked with the detection of these nuances. To synthesize multiple anomaly indicators, we propose to use the Boltzmann machine as a promising method. The theory of the Boltzmann machine and an illustrative example are explained in the appendix, Synthesis of Comprehensive Index from Individual Anomaly Features. This machine produces an anomaly synthesis indicator, denoted as A, by integrating various anomaly indicators. Concurrently, a quantum Boltzmann machine-based AI is developed, generating a comprehensive anomaly indicator, denoted as B. The success or failure of the latter AI hinges on advancements in quantum computation algorithms.

(3) *Step 3*: upon obtaining an anomaly prediction, the transaction network's graph Laplacian eigenvectors and the

Ricci curvature of each edge, determined through the optimal transport problem, are leveraged to identify the node accountable for the anomaly. Step 3 in Fig. 2 illustrates this process on the right-hand side.

5 Facilitation AI

Another use of AI in this research is to proactively support traders in making appropriate trading decisions in response to anomalies predicted in the previous section. A suitable framework for this purpose could be built as a facilitation AI platform relying on intelligent and autonomous conversational agents, known also as conversational multiagent systems (Hadfi and Ito 2022). Facilitation AI is often referred to as conversational AI because it relies on conversational agents. The first purpose of these agents is to alert DAO members in the event of price spikes or other anomalies. Recall that the "typical" DAO members include traders, system developers, and operators. In our AI-driven platform, the traders will interact with artificial agents and use their facilitative capacities to work towards appropriate trading decisions.

One technical challenge here is to integrate the conversational agents with agent-assisted wallets. Such intelligent wallets will manage the private keys of the traders and protect their assets. Moreover, the integrated agents will autonomously manage sending, receiving, and spending cryptocurrencies in response to the decisions of the traders. The facilitation AI framework is implemented with interactive interfaces for the dialogues with the traders on the underlying Social Networking Services (SNSs).

The proposed framework is illustrated in Fig. 3 and elucidates a synergistic integration of predictive AI, facilitation AI, and DAOs within an AI-driven, blockchain-centric intelligent market. This framework could extend the fintech domain to civic engagement and broader global ecosystems (Viano et al. 2023). In practice, predictive AI employs mathematical models to analyze the transactional data of crypto-assets, scanning for anomalous patterns of events across various market conditions (1). Upon the detection of anomalies, notifications are dispatched to both the DAO and the facilitation AI (2). Once validated through a blockchain oracle that acts as an event filter (3), such anomalies are then relayed to the DAO (Eskandari et al. 2021).

The facilitation AI is pivotal in propagating notifications throughout a network of traders equipped with conversational agents and digital wallets (4). Detected anomalies and events are articulated textually, triggering discussions between the traders, potentially inciting collective decision-making or prompting actions executed by their surrogate agents (5). Such actions manifest on the off-chain social network as deliberation (Hadfi and Ito 2022); on the digital wallet as transactions; and on the DAO as trades controlled by smart contracts. The DAO is intended to orchestrate such actions in response to incoming market anomalies. This adaptive response (6) is amplified by reinforcement learning, connecting the agents and wallets back to the DAO. Within this feedback loop (6), the DAO's function relies on the exchange of certificates and smart contracts that behave

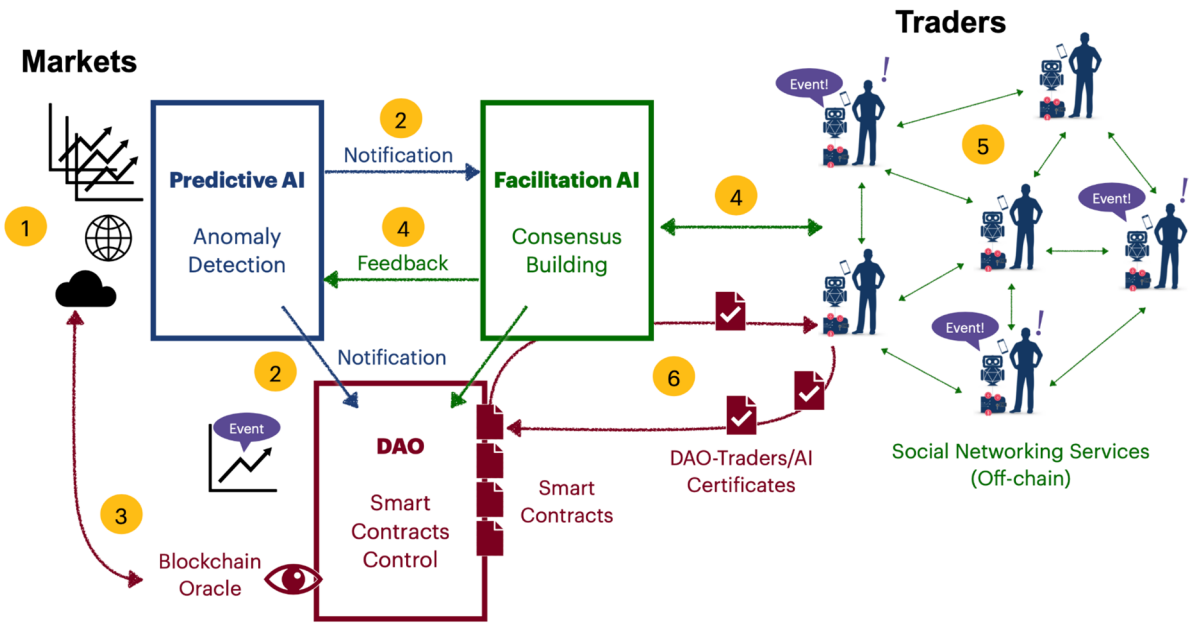

Fig. 3 Integration of facilitation AI with predictive AI and DAO within an intelligent market

adaptively through constant learning and adaptive control that enforces the decisions of the traders on the DAO level.

The framework illustrated in Fig. 3 could be developed and deployed on Amazon Web Services (AWS). If the number of users increases, the number of AWS nodes can be scaled accordingly. In principle, there is no upper limit to the number of DAO members. This scalability is facilitated by the framework's design, which follows the Model-view-controller design pattern (Krasner and Pope 1988) where the predictive AI acts as the "Model" with its mathematical predictive logic attempting to model various aspects of the markets (anomalies, uncertainty, volatility, risks, etc.). The facilitation AI presents the "View", processing, and presenting the events linguistically to the traders. The DAO emulates the "Controller", certifying and orchestrating the interactions between the "Model" and "View" via smart contracts embedded within the blockchain infrastructure.

6 Smart contract control

In the prediction-consensus-control process orchestrated by stakeholders within DAOs, illustrated in Fig. 1, smart contracts seamlessly integrate external information via an oracle, enacting effective control measures in response to predicted anomalies. The smart contracts of the DAOs offer an alternative to conventional policy implementation by administrative bodies. Traditional policy implementation is often slow, requiring patience from citizens. In contrast, smart contracts execute instantly upon meeting predefined conditions, eliminating the need for prolonged waiting periods. The substitution of policy implementation with smart contracts can thus be viewed as a paradigm shift toward society being governed by these automated contractual agreements.

It is important to note that the facilitation provided by software agents may prove inadequate due to insufficient accuracy in preemptively detecting signs of anomalies, potentially leading to unintended losses if acted upon. Consequently, there arises a need for an insurance mechanism, implemented through smart contracts, to mitigate such losses. However, the occurrence of widespread losses among stakeholders simultaneously, resulting from spikes and crashes in crypto-asset prices induced by anomalies, raises doubts about the feasibility of insurance coverage. The concern emerges from the likelihood of insurance underwriters abstaining from participation in such scenarios. Yet, the integration of financial instruments like options offers a viable solution (Rahman et al. 2022). Within the context of options, insurance functionalities can be effectively implemented, as there exists a continuous market of investors willing to engage in transactions involving risky instruments. To safeguard against unintended losses, users seeking protection can acquire "put options" with crypto-assets as the underlying asset from external investors, thus fortifying their assets against adverse market movements. The market size for options for major crypto assets is growing quickly. The monthly trading volume for Ethereum futures options on the Chicago Commodity Exchange (CME) reached $1.26 billion in May 2024. For minor crypto assets, there is concern that market liquidity is small.

Consider, for instance, a put option entitling the holder to sell one unit of a crypto-asset to the investor at a strike price of $220 at time t = T. This put option is acquired by a user from the investor through the DAO at a price denoted as π. Let us assume the user possesses 0.6 units of the crypto-asset from t = 0 to t = T. At t = 0, the price of the crypto-asset stands at $300, and by t = T, it is projected to either surge to $500 or plummet to $100. Given the initial ownership of 0.6 units of the crypto-asset at t = 0, the user's total asset value amounts to $300 × 0.6 = $180. At t = T, if the crypto-asset plummets to $100, with the price of the crypto-asset serving as the oracle, the put option is automatically executed by the smart contract. If the put option is acquired by paying the price π, the total asset value remains at $180, safeguarding against losses in the event of a price decline. Because the value of crypto-assets owned is $100 × 0.6 =$60 and the value of the put options purchased is $220 −$100 =$120. However, the precise pricing of such put options is pivotal for trading decisions, necessitating the development of a stochastic process model to accurately compute option prices utilizing realistic crypto-asset price time series.

Moreover, an additional criterion for the execution of the option may entail ensuring that the buyer of the put option is not engaged in any anomalous transactions, in conjunction with verifying the price of the crypto-asset through the oracle. Such measures aim to enhance the security for investors selling put options. Transparency in financial transactions and traceability play pivotal roles in deterring potential misuse of options for the advantage of specific entities or cohorts. By scrutinizing transactions logged on the blockchain pertaining to users who have procured options, it becomes feasible to validate that said users have neither transferred funds to nor received funds from addresses associated with anomalous activities. Similar concepts are discussed for escrow services in crypto-asset transactions.

Figure 4 illustrates how different stakeholders use the DAO platform. Governments publish their regulations, policies, and smart contract templates to guarantee non-anomalous transactions on the blockchain. Traders could then perform all transactions via intelligent software agents integrated with wallets that also guarantee secure transactions. The agents can also rely on predictions from the predictive AI on the characteristics of traders to facilitate transactions with the DAO. All these interactions between traders and agents are recorded as immutable evidence on the blockchain. Note that blockchain technology has an issue

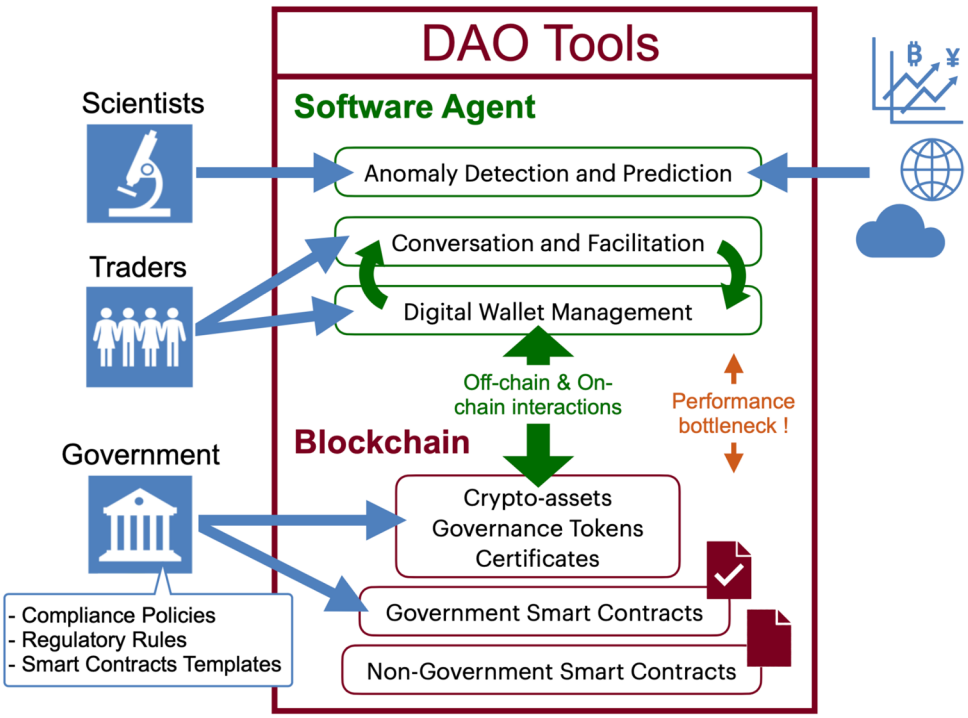

Fig. 4 Software agents and blockchain integration for different stakeholders

in real-time transaction processing. As depicted in Fig. 4, the transaction processing performance of recording such evidence could result in a bottleneck in the framework. It is, therefore, necessary to build robust mathematical models of the DAO to evaluate performances theoretically. Developing a DAO simulator and evaluating its performance in realistic scenarios is also necessary. Based on the theoretical evaluations and simulation results, it is possible to examine the improvements in real-time transaction processing and the blockchain's operations in ranges that do not exceed the performance limit.

The Government DAO in Fig. 4 is basically the result of government action that makes regulation machine-readable and executable. Traditionally, the government is bound by local law in its decisions, and a parallel decision-making mechanism could be problematic. In addition to making regulations machine-readable and executable, they attempt to go beyond the simple automation of existing governmental functions (introduction of decentralized, blockchain-based governance that operates transparently and autonomously within predefined rules). While it is true that government actions are bound by local law, the integration of DAOs offers a novel paradigm where laws can be encoded as smart contracts, allowing for more efficient and automated compliance. This adoption can reduce the bureaucratic delays in traditional governance and enhance accountability. We note that the interplay between national regulations and global protocols is crucial, and interoperability frameworks must be designed to ensure that DAOs can function across jurisdictions without creating conflicts.

The smart contract control might be limited by the fact that malicious actors can deduce the policy of the system and devise strategies to game it. Similar issues were discussed for escrow services in crypto asset transactions (Siebenbrunner and Taudes 2024). For this reason, actual implementation needs to consider having a human in the loop in case of errors and leave the reaction at least partly to a DAO. The human-in-the-loop aspect is indeed addressed in our architecture, as illustrated in Fig. 3. Specifically, human validators can interact with the DAOs through AI agents, ensuring that critical decisions are not left entirely to smart contracts. This design is inspired by the concept of "reinforcement learning from human feedback (RLHF)", where human supervision is integrated to mitigate risks in Machine Learning systems, such as system gaming by malicious actors or errors in automated responses (Dai et al. 2023). By combining human judgment with the speed and efficiency of smart contracts, our conception seeks to balance rapid reactions with trustworthy supervision, reducing vulnerabilities while maintaining agility.

7 Summary

This proposal sets the ground for a framework that supports traders in their economic activities within cyberspace. By detecting and alerting traders to criminal activities and anomalies, it is possible to enhance public trust in both crypto-assets and the underlying blockchain technology. The realization of DAOs equipped with the capability to detect

and notify traders of such anomalies stands as a viable, long-term goal. To operationalize this goal, we propose a framework that integrates predictive AI, facilitation AI, and DAOs within an AI-driven, blockchain-centric intelligent market. The resulting DAOs are envisioned to complement the decision-making processes of traditional joint stock companies, fostering a global economic landscape where democratic participation is open to all individuals. Moreover, users seeking to mitigate unforeseen losses can safeguard their assets by engaging in put options with crypto-assets as the underlying asset, procured from external investors. The execution of these options will be seamlessly managed through smart contracts, ensuring asset protection for users. By bridging economic activities across physical and cyberspace domains, a transformative shift in values and lifestyles is anticipated, nurturing the emergence of a new societal paradigm focused on wellbeing and sustainability, diverging from conventional GDP-driven economic growth. Such a proposal comes with a number of challenges including the need for reliable anomaly detection techniques, trustworthy conversational AI, secure and safe agent wallets, and the existence of information bottlenecks on the blockchain level.

8 Appendix

9 Appendix: synthesis of comprehensive index from individual anomaly features

From individual anomaly features evaluated from individual analyses, a Boltzmann machine is used to synthesize a comprehensive index. We use the restricted Boltzmann machine (RBM), in which there are interactions only between visible and hidden variables and no interactions among visible variables and among hidden variables. The RBM is equivalent to a bipartite graph if the variable and interaction are expressed as node and link. The Boltzmann machine is a neural network that learns parameters that reproduce a given binary variable. In contrast, the Ising model is a phase transition model that computes a binary variable with given parameters. In this sense, The Boltzmann machine is an inverse problem of the Ising model. In this appendix, we explain a test calculation of the RBM with a 6-dimensional Binarized feature vector time series (N=6) and 2 hidden variables (M =2).

We define N dimensional visible variable $\boldsymbol{v} = \left(v_1, v_2, \cdots, v_N\right)$ and M dimensional hidden variable $\boldsymbol{h} = \left(h_1, h_2, \cdots, h_M\right)$. The energy of the system or Hamiltonian of the RBM is written as

$$E(\boldsymbol{v}, \boldsymbol{h}) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_{ij} v_i W_{ij} h_j \tag{A.1}$$

where parameters are $\theta = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{W})$ and the interactions between visible variables and among hidden variables are ignored. The joint probability of the system's state exhibits the Boltzmann distribution:

$$P(\boldsymbol{v}, \boldsymbol{h}) = \frac{\exp(-E(\boldsymbol{v}, \boldsymbol{h}))}{\sum_{v,h} \exp(-E(\boldsymbol{v}, \boldsymbol{h}))} \tag{A.2}$$

The parameters are learned by a stochastic sampling method called the contrastive divergence method:

$$\theta^{new} = \theta^{old} + \varepsilon \left[\sum_{k=1}^{N} \frac{\partial\left(-E\left(v_k, h_k\right)\right)}{\partial\theta}_{P(h_k|v_k)} - N \frac{\partial(-E(v,h))}{\partial\theta}_{P(\boldsymbol{v},\boldsymbol{h})} \right] \tag{A.3}$$

The model performance is evaluated using the F1 score:

$$F1 = \frac{2\mathrm{TP}}{2TP + FP + FN} \tag{A.4}$$

where TP: model prediction is 1 and real data is 1, FP: model prediction is 1 and real data is 0, TN: model prediction is 0 and real data is 0, and FN: model prediction is 0 and real data is 1.

In the test calculation, we first conduct the learning of model parameters. The data reconstruction is to compute the hidden variables from the input visible variables and output the computation of the visible variables from the computed hidden variables. The input visible variables are N-dimensional data (N =6), with two to three 1 s, where we assume fewer 1 s in normal periods, as shown in Fig. 5. We learn the model parameters in a normal period. Learning of model parameters was done using the contrastive divergence method in Eq. (A.3). Figure 6 shows the reconstructed visible variables match the input visible variables with F1 = 0.833. This means that the learning results are good.

Next, we carried out anomaly detection using parameters learned from normal period data. Input 6-dimensional test data are with 4 to 5 1 s, as shown in Fig. 7, where we assume that there are more 1 s for the anomaly event. Figure 8 shows the reconstructed visible variables that do not match the input visible variables (F1 = 0.0). This means

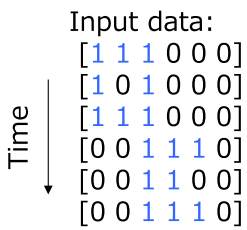

Fig. 5 The input visible variables in the normal period

Time	Reconst. data:	Visible energy:	Hidden energy:	Match:
	[1 1 1 0 0 0]	-22.272	-5.006	True
	[1 0 1 0 0 0]	-24.151	-0.017	True
	[1 1 1 0 0 0]	-22.272	-5.006	True
	[0 0 1 1 1 0]	-21.916	-13.736	True
	[0 0 1 1 **1** 0]	-21.916	-7.531	**False**
	[0 0 1 1 1 0]	-21.916	-13.736	True

Fig. 6 The reconstructed visible variables in learning model parameters

Fig. 7 The input visible variables in the anomaly detection

Time	Input data:
	[1 1 1 0 1 1]
	[1 0 1 1 0 1]
	[1 1 1 1 1 0]

Time	Reconst. data:	Visible energy:	Hidden energy:	Match:
	[1 1 1 0 **0 0**]	-22.272	-6.323	**False**
	[**0** 0 1 1 0 **0**]	-21.916	-0.805	**False**
	[1 1 1 **0 0** 0]	-22.272	-1.020	**False**

Fig. 8 The reconstructed visible variables in the anomaly detection

that the reconstruction of input visible variables is poor for the anomaly event. We did not observe much difference in the system's energy in the normal events and the anomaly events. Therefore, energy cannot be used to determine anomalies. However, we emphasize that a series of false reconstructions of the input visible variables can be regarded as an anomaly. Thus, we use a series of false of the reconstruction as a comprehensive anomaly index.

Funding Silicon Valley Community Foundation, 2022-247584, Yuichi Ikeda.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Ahmed M, Mahmood AN, Islam MR (2016) A survey of anomaly detection techniques in financial domain. Futur Gener Comput Syst 55:278–288

Altaleb H, Zoltán R (2022) Decentralized autonomous organizations review, importance, and applications, 2022 IEEE 26th International Conference on Intelligent Engineering Systems (INES), Georgioupolis Chania, Greece, pp. 000121–000126, https://doi.org/10.1109/INES56734.2022.9922656

Aoyama H, Fujiwara Y, Hidaka Y, Ikeda Y (2022) Cryptoasset networks: flows and regular players in Bitcoin and XRP. PLoS One 17(8):e0273068. https://doi.org/10.1371/journal.pone.0273068

Berman S, Snegovaya M (2019) Populism and the decline of social democracy. J Democracy 30:5

Buterin V (2024) The promise and challenges of crypto + AI applications, https://vitalik.eth.limo/general/2024/01/30/cryptoai.html

Chakraborty A, Hatsuda T, Ikeda Y (2023) Projecting XRP price burst by correlation tensor spectra of transaction networks. Sci Rep 13:4718. https://doi.org/10.1038/s41598-023-31881-5

Chakraborty A, Hatsuda T, Ikeda Y (2024) Dynamic relationship between the XRP price and correlation tensor spectra of transaction networks. Physica A 639(1):129686. https://doi.org/10.1016/j.physa.2024.129686

Dai J, Pan X, Sun R, Ji J, Xu X, Liu M, Wang Y, Yang Y (2023) Safe rlhf: Safe reinforcement learning from human feedback. arXiv preprint arXiv:2310.12773

Eskandari S, Salehi M, Gu WC, Clark J. Sok: Oracles from the ground truth to market manipulation. Proceedings of the 3rd ACM Conference on Advances in Financial Technologies. 2021. https://doi.org/10.1145/3479722.3480994

Hadfi R, Okuhara S, Haqbeen J, Sahab S, Ohnuma S, Ito T (2023) Conversational agents enhance women's contribution in online debates. Sci Rep 13:14534. https://doi.org/10.1038/s41598-023-41703-3

Hadfi R, Ito T (2022) Augmented democratic deliberation: can conversational agents boost deliberation in social media? Proceedings of the 21st International Conference on Autonomous Agents and Multiagent Systems

Hadfi R, Ito T (2023) Computational social choice competition: overview. The 37th Annual Conference of the Japanese Society for Artificial Intelligence, pages 2F6GS501–2F6GS501

Hilal W, Gadsden SA, Yawney J (2022) Financial fraud: a review of anomaly detection techniques and recent advances. Expert Syst Appl 193:116429. https://doi.org/10.1016/j.eswa.2021.116429

Ikeda Y, Chakraborty A (2023) Hodge decomposition of the remittance network on the XRP ledger in the Price Hike of January 2018, JPS Conf. Proc. 40, 011004, Proceedings of Blockchain Kaigi 2022 (BCK22) https://doi.org/10.7566/JPSCP.40.011004

Ikeda Y (2022) Characterization of XRP crypto-asset transactions from networks scientific approach. In: Aruka Y (eds) Digital designs for money, markets, and Social Dilemmas". Evolutionary Economics and Social Complexity Science, vol 28. Springer, Singapore. https://doi.org/10.1007/978-981-19-0937-5_8

Yuichi Ikeda et al (2024) Verification of elemental technologies for anomaly detection in crypto asset transactions, RIETI Discussion Paper Series 24-E-085, December 2024, the Research Institute of Economy, Trade and Industry (RIETI)

Ito T, Hadfi R, Suzuki S (2022) An agent that facilitates crowd discussion. Group Decis Negot 31:621–647. https://doi.org/10.1007/s10726-021-09765-8

Krasner GE, Pope ST (1988) A description of the model-view-controller user interface paradigm in the smalltalk-80 system. J Object Orient Programm 1(3):26–49

Rahman A, Shi V, Ding M, Choi E (2022) Systematization of knowledge: synthetic assets, derivatives, and on-chain portfolio management", arXiv:2209.09958 [q-fin.GN]

Reurink A (2019) Financial fraud: a literature review. In Contemporary Topics in Finance (eds I. Claus and L. Krippner). https://doi.org/10.1002/9781119565178.ch4

Siebenbrunner C, Taudes A (2024) Why CBDCs will likely not support full smart contracts. Working Paper Series, Research Institute for Cryptoeconomics, WU Vienna. https://doi.org/10.57938/aaebf4b1-6fe8-490d-b40f-84328df75adf

Takefuji Y (2023) Case report on enormous economic losses caused by fraud from Japan to the world. J Econ Criminol 1:100003. https://doi.org/10.1016/j.jeconc.2023.100003

Tamai S, Kasahara S (2024) DAO voting mechanism resistant to whale and collusion problems. Front Blockchain. https://doi.org/10.3389/fbloc.2024.1405516

Thudumu S, Branch P, Jin J, Singh JJ (2020) A comprehensive survey of anomaly detection techniques for high dimensional big data. J Big Data. https://doi.org/10.1186/s40537-020-00320-x

Valla R, Mozharovskyi P, d'Alché-Buc F (2023) Anomaly component analysis, arXiv:2312.16139 [stat.ME]

Viano C, Avanzo S, Boella G, Schifanella C, Giorgino V (2023) Civic Blockchain: making blockchains accessible for social collaborative economies. J Respons Technol 15:100066. https://doi.org/10.1016/j.jrt.2023.100066

Wattenhofer R, Feichtinger R, Fritsch R, Heimbach L, Vonlanthen Y (2024) SoK: attacks on DAOs, 4th Workshop on Decentralized Finance (DeFi 2024), Willemstad, Curaçao

Yano M, Dai C, Masuda K, Kishimoto Y (2020) Creation of blockchain and a new ecosystem. Blockchain and Cryptocurrency 1