Glovento Journal of Integrated Studies (GJIS) | ISSN: 3117-3314
Volume 1 (2025) | Article 30

Sustainable Development of Digital Financial Infrastructure: Evaluating Data Quality and Information Fairness in Cryptocurrency Exchanges

[1,*]Jun-Kuan Tsao, [2]Chiang-Yu Cheng

[1,2]Department of Data Science, Soochow University, Taipei, Taiwan
Email: [1]junkuancao35@gmail.com, [2]taicanccy@gmail.com
*Corresponding Author: Jun-Kuan Tsao (Email: junkuancao35@gmail.com)

Abstract- The quality of digital financial infrastructure directly impacts the fairness of information access for market participants, which is fundamental to the sustainable development of the entire financial ecosystem. As cryptocurrency markets operate globally 24/7 without interruption, ensuring equitable access to high-quality data is essential for promoting responsible investment and reducing market inequality. This study establishes a standardized data quality assessment framework for three major cryptocurrency exchanges to explore how data quality differences affect information fairness. Through systematic data collection conducted every 60 seconds for eight consecutive hours, 463 valid Bitcoin price data samples were successfully obtained, covering major Asian and European trading sessions. The research constructs a comprehensive Data Quality Index integrating three core dimensions of consistency, freshness, and availability metrics. The consistency component measures price synchronization between platforms by calculating relative price differences with a coefficient of 50,000 to capture micro-level variations. The freshness indicator evaluates information timeliness by measuring delays between API request time and exchange data timestamps using a stringent 0.2-second threshold. The availability metric assesses service reliability through success rate calculations. The final composite index employs weighted averaging with 40% for consistency, 40% for freshness, and 20% for availability. Results reveal that Binance ranks first with 92.50 points, followed by Bitget at 88.61 points and MEXC at 85.36 points, with the 7.14-point gap directly impacting information fairness. This study provides quantitative assessment tools for building sustainable digital financial ecosystems, supporting UN Sustainable Development Goals 9.1 and 10.5.

Keywords: Digital Financial Infrastructure, Data Quality Assessment, Information Fairness, Cryptocurrency Exchanges, Bitcoin.

I. INTRODUCTION

Digital financial infrastructure has become a crucial pillar of modern economic systems, with its service quality directly affecting the decision-making efficiency of market participants and the stability of financial systems(Ozili, 2018). Cryptocurrency exchanges, as core platforms for digital asset trading, undertake the critical functions of price discovery, liquidity provision, and information transmission (Brandvold et al., 2015). According to the latest market data, the global cryptocurrency market capitalization has exceeded 4 trillion USD, currently standing at 4.05 trillion USD, demonstrating its growing importance in the global financial system (BlockBeats, 2025). As market scale continues to expand, cryptocurrency trading has become an indispensable component of the global financial system(Corbet et al., 2019).

However, information asymmetry problems that exist in financial markets are equally present in cryptocurrency trading environments, where data quality differences across exchanges may impact market efficiency(Admati& Pfleiderer, 1988). The World Bank emphasizes that high-quality financial infrastructure is a key element for achieving inclusive finance and reducing inequality(World Bank, 2022). Data quality theory indicates that information users' demands for data extend beyond accuracy to include multiple dimensions such as consistency, timeliness, and availability (Wang & Strong, 1996). The high volatility characteristics of cryptocurrency markets make data timeliness even more critical (Hafner, 2018). For high-frequency traders, data delays can lead to trading disadvantages; for general investors, inconsistent price information may result in erroneous investment decisions.

These information quality difference issues are related to the United Nations Sustainable Development Goals of reducing inequalities (SDG 10) and industry innovation and infrastructure development (SDG 9) (United Nations General Assembly, 2015). The development of financial technology provides a technical foundation for establishing a more equitable and transparent financial environment (Arner et al., 2015). Therefore, establishing an objective and standardized data quality assessment framework not only helps improve market efficiency but is also a necessary measure for promoting the sustainable development of digital financial infrastructure.

This study aims to establish a systematic data quality assessment framework for cryptocurrency exchanges, quantitatively analyze performance differences in data services among major exchanges, and explore the impact of these differences on information fairness. The research establishes a standardized data quality assessment framework targeting three major cryptocurrency exchanges (Binance, MEXC, and Bitget), constructing a comprehensive Data Quality Index (DQI) that encompasses consistency indicators, data freshness indicators, and availability indicators through systematic data collection conducted every 60 seconds over eight consecutive hours. The research results will provide quantitative assessment tools for building sustainable digital financial ecosystems, supporting the realization of UN Sustainable Development Goals 9.1 (reliable infrastructure development) and 10.5 (improved financial market regulation).

II. LITERATURE REVIEW AND DISCUSSION

The theoretical foundation of information fairness can be traced back to classical market microstructure theory. Kyle (1985) established a theoretical framework linking information asymmetry and market efficiency in his seminal research, indicating that informed traders exploit information advantages to obtain excess profits, which damages overall market efficiency and fairness. The intraday trading pattern theory by Admati and Pfleiderer (1988) demonstrates that the degree of competition among informed traders affects the welfare of liquidity traders, and when the information environment promotes more sufficient competition, it can improve welfare distribution among market participants. Diamond and Verrecchia (1991) showed that appropriate information disclosure can reduce corporate capital costs and enhance market liquidity through reducing information asymmetry.

In terms of data quality assessment methodology, Wang and Strong (1996) established a multi-dimensional information quality framework, categorizing data quality into four categories: intrinsic, contextual, representational, and accessibility. This framework emphasizes that effective data quality assessment must consider objective indicators across multiple dimensions, as single-dimension assessments often cannot comprehensively reflect the true state of data quality. This research provides important theoretical foundations for the indicator design in our study, particularly in conceptualizing consistency and availability indicators.

Research on information transmission mechanisms in cryptocurrency markets has gradually increased in recent years. The study by Brandvold et al. (2015) represents an important milestone in this field, discovering obvious price discovery hierarchical structures among Bitcoin exchanges, where large exchanges often play leadership roles in the price discovery process, while smaller exchanges tend to follow price movements. This finding provides theoretical basis for our study's selection of different-sized exchanges for comparative analysis. Schwenkler et al. (2025) found that among multiple cryptocurrency data providers from 2018-2024, approximately 21% of coin IDs were replaced without disclosure, and 70% of daily closing prices deviated more than 5% from peer median values, highlighting the criticality of standardized data for market transparency and efficiency.

Data timeliness holds special importance in cryptocurrency markets. Alexander (2025), through empirical analysis of high-frequency order flows at multiple institutional-grade exchanges, found that microsecond-level delays can significantly affect arbitrage returns and liquidity estimation, thereby changing optimal execution paths and risk control decisions, highlighting the critical role of data freshness in cryptocurrency market analysis. The study documented latency variations of 30-40 milliseconds between exchanges and showed that high-frequency traders with lower latency could capture profit margins of 15-20 basis points per trade relative to competitors with higher latency. This finding directly supports the rationality of our study treating data freshness as a core assessment dimension.

Despite rich achievements in existing literature regarding information fairness, data quality assessment, and cryptocurrency market research, obvious research gaps remain. First, most existing research focuses on data quality issues in traditional financial institutions, with relatively few systematic evaluation studies specifically targeting cryptocurrency exchange data quality. Second, cryptocurrency market research mostly focuses on price prediction, risk management, or regulatory policies, with less exploration of exchange service quality differences and their impact on market participants from an information fairness perspective. Finally, research frameworks combining data quality assessment with sustainable development concepts have not been sufficiently developed, lacking methodologies for translating technical indicators into social benefit assessments. This study attempts to fill these research gaps, providing new research perspectives and empirical evidence for the sustainable development of digital financial infrastructure.

III. RESEARCH METHOD

3.1. Research Design
This study employs a quantitative empirical analysis approach, collecting actual data through standardized API interfaces and establishing a multi-dimensional evaluation framework to compare the data quality performance of cryptocurrency exchanges. The detailed research process is shown in Figure 1.

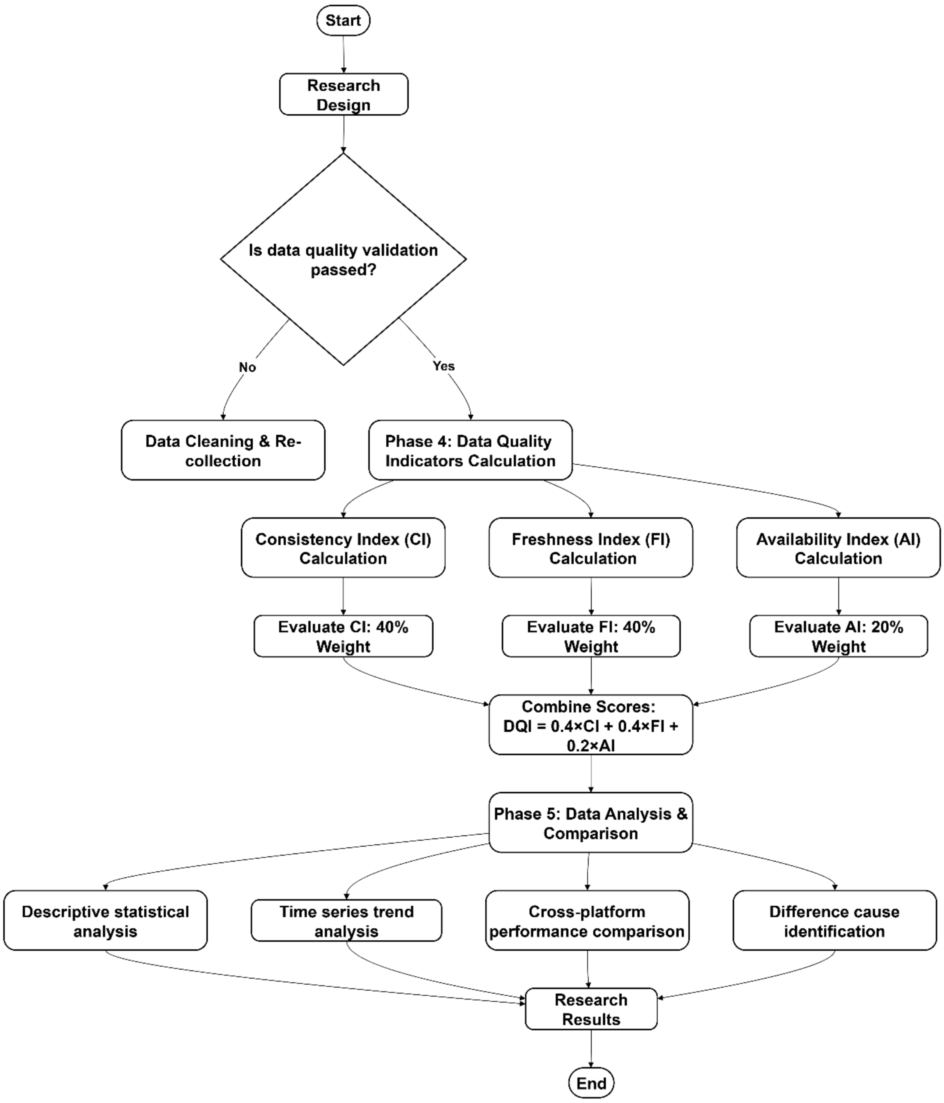

Figure 1: Data Quality Assessment Framework for Cryptocurrency Exchanges

3.2. Research Objects and Data Collection
3.2.1 Selection of Research Objects

This study selects three representative cryptocurrency exchanges. Binance serves as the world's largest exchange, representing major international platforms; MEXC is a medium-sized regional exchange with significant influence in the Asian market; Bitget is an emerging medium-sized exchange representing the characteristics of new-generation trading platforms.

BTC/USDT is selected as the target trading pair because Bitcoin, as the benchmark asset in the cryptocurrency market, possesses the highest market attention and liquidity, thus providing good market representativeness. BTC/USDT is a standard trading pair across all mainstream exchanges, ensuring consistency for cross-platform comparisons.

3.2.2Data Collection Method
Data collection period was set from 11:14 to 19:14 on September 6, 2025, for continuous 8 hours, with a collection frequency of once every 60 seconds, obtaining 463 valid samples with a theoretical sample size

of 480. API endpoints used include Binance's /api/v3/ticker/price, MEXC's /api/v3/ticker/price, and Bitget's /api/spot/v1/market/ticker.

The data collection system was implemented using Python 3.8+ with a multi-threaded architecture to ensure simultaneous API requests to all three exchanges, minimizing time differences between different platforms and improving the reliability of comparative analysis. The core implementation includes request handling modules, timestamp recording functions, data validation mechanisms, and real-time storage capabilities. The system employs the requests library for HTTP communication, threading module for concurrent processing, and json library for data parsing. Each API call captures both request timestamp and exchange-provided timestamp, enabling precise delay calculation for freshness evaluation.

To ensure analysis accuracy, the system records dual timestamp information: $T_{request}$ representing API request sending time and $T_{exchange}$ representing exchange data timestamp. This design enables accurate calculation of data delay time and evaluation of data freshness performance across exchanges.
The collected dataset comprises 463 valid samples with 15 attributes per sample, including timestamp information, success indicators, price data, and platform-specific metrics for each exchange. Data quality validation procedures were implemented to ensure price reasonableness and timestamp consistency. The raw dataset and processed analysis results are stored in both JSON and CSV formats for comprehensive accessibility and reproducibility.

All source code for the data collection system, analysis algorithms, and evaluation framework implementations, along with the complete dataset obtained from this study, are publicly available (Tsao, 2025). This repository includes the multi-threaded data collection scripts, real-time analysis tools, visualization components, and comprehensive documentation to ensure research reproducibility and facilitate future extensions of this work.

3.3. Evaluation Indicator Design

3.3.1Consistency Index (CI)
Evaluates the degree of price synchronization between different exchanges at the same time point:
$$CI_i = \max\left(0, 100 - \frac{|P_i - \bar{P}_{-i}|}{\bar{P}_{-i}} \times 50000\right) \quad (1)$$

Where P_i is the price of exchange i, and $\bar{P}_{-i}$ is the average price of other exchanges. The consistency index reflects the fairness of price information transmission and market efficiency. The coefficient of 50,000 is designed to capture micro-level price differences of 0.001%, ensuring sensitivity to changes in market microstructure. When prices across different exchanges are completely consistent at the same time point, the indicator scores 100 points; the greater the price difference, the lower the score. This indicator is significant for evaluating arbitrage opportunities and the degree of information asymmetry. Platforms with higher price consistency can provide users with a fairer trading environment and reduce trading disadvantages caused by price differences.

3.3.2Freshness Index (FI)
Evaluates information timeliness by first calculating delay time:
$$\Delta t = \left|T_{request} - T_{exchange}\right| \quad (2)$$
Then adopting a segmented scoring mechanism. When delay time is within 0.2 seconds, a full score of 100 points is awarded; for delays of 0.2-0.5 seconds, 0.167 points are deducted for every 0.01-second increase; for delays of 0.5-1.0 seconds, 0.3 points are deducted for every 0.01-second increase; for delays of 1.0-2.0 seconds, 0.3 points are deducted for every 0.01-second increase; for delays of 2.0-5.0 seconds, 0.1 points are deducted for every 0.01-second increase; for delays exceeding 5.0 seconds, 0.08 points are deducted for every 0.01-second increase, with a minimum score of 0 points. This segmented design reflects the differential impact of various delay levels on trading decisions in high-frequency trading environments. The 0.2-second threshold is based on Alexander (2025)'s research demonstrating the critical influence of microsecond-level delays on trading outcomes.

3.3.3 Availability Index (AI)
Measures service stability:
$$AI = \frac{N_{success}}{N_{total}} \times 100 \quad (3)$$

Where $N_{success}$ is the number of successful requests and N_{total} is the total number of requests. The availability index reflects the stability and reliability of exchange API services, serving as an important dimension for evaluating basic data service quality. In modern financial trading environments, system availability directly affects whether users can obtain market information in a timely manner and execute trading decisions. 100% availability indicates that all API requests received successful responses throughout the entire observation period, while lower availability may prevent users from obtaining necessary market information at critical moments, affecting trading fairness.

3.3.4 Composite Data Quality Index (DQI)
Adopts weighted averaging:

$$DQI = 0.4 \times CI + 0.4 \times FI + 0.2 \times AI \quad (4)$$

The Composite Data Quality Index integrates three core dimensions, providing an overall assessment of exchange data services. The weight allocation is based on market microstructure theory, with consistency and timeliness as the two most critical factors affecting trading fairness, each accounting for 40% weight, reflecting their importance in high-frequency trading and arbitrage environments. The availability index accounts for 20% weight because differences in basic service stability among mainstream exchanges are relatively small, but it remains a fundamental requirement for ensuring users can access data services. DQI scores range from 0-100 points, with higher scores indicating better data quality and the ability to provide market participants with a fairer, more timely, and reliable information environment. This composite index not only facilitates comparative analysis between different exchanges but also provides quantitative reference criteria for regulatory agencies in establishing industry standards and for users in selecting trading platforms.

IV. RESEARCH RESULTS

4.1. Basic Statistical Results
The research successfully collected 463 valid data samples over a continuous 8-hour observation period from 11:14 to 19:14 on September 6, 2025, achieving 96.46% of the theoretical sample size of 480. All three exchanges achieved 100% API success rates, demonstrating stable basic service performance.
During the observation period, Bitcoin exhibited relatively stable price movements with an average price of $110,742.36 and standard deviation of $108.23. The price range spanned from $110,580 to $111,100, with intraday volatility of approximately $520. The observation period covered both Asian trading session (226 samples) and European trading session (237 samples), ensuring balanced temporal distribution for cross-session analysis.

4.2. Composite Data Quality Index Assessment Results Overview
Based on the established multi-dimensional evaluation framework, the Composite Data Quality Index (DQI) assessment results for the three exchanges demonstrate clear hierarchical differences. The comprehensive evaluation results are presented in Table 1.

Table 1: Data Quality Index Assessment Results for Three Cryptocurrency Exchanges

	Composite DQI	Consistency Index	Freshness Index	Availability Index	Average Delay (sec)
Binance	92.5	97.72	83.53	100	0.878
Bitget	88.61	93.58	77.95	100	1.066
MEXC	85.36	94.36	69.06	100	1.362

Binance ranks first with 92.50 points, establishing a clear leading advantage; Bitget ranks second with 88.61 points, showing stable and balanced performance; MEXC ranks third with 85.36 points, primarily limited by data freshness performance.

The DQI score gap among the three exchanges spans 7.14 points, with the difference between the highest and lowest scores sufficient to substantially impact the fairness of information access for market participants. According to the evaluation framework's design standards, a gap exceeding 7 points indicates significant differences in data service quality among platforms, potentially affecting users' trading decision efficiency and market participation fairness.

Glovento Journal of Integrated Studies (GJIS) | ISSN: 3117-3314
Volume 1 (2025) | Article 30

Figure 2: BTC Price Movements by Exchange

Analysis of the Bitcoin price movements across both trading sessions (Figure 2) reveals high price consistency among all three exchanges throughout the 8-hour observation period. The price trajectory shows an initial decline from approximately $111,300 at 12:00 to around $110,700 by 14:00, followed by a gradual recovery to $110,800 during the European session opening around 17:00, before stabilizing around $110,750 toward the end of the observation period. Notably, all exchanges maintained remarkably similar price levels across both Asian and European sessions, with the maximum price deviation between platforms rarely exceeding 0.002%, confirming the high consistency scores observed in the quantitative assessment.

4.3 Detailed Analysis of Dimensional Performance

4.3.1 Price Consistency Performance Analysis

All three exchanges demonstrated excellent performance in price consistency, with scores concentrated in the high-level range of 93-98 points. Binance leads with 97.72 points, significantly outperforming Bitget's 93.58 points and MEXC's 94.36 points. Interestingly, MEXC slightly outperforms Bitget in consistency, showing that despite its lower overall DQI ranking, it maintains good performance in price synchronization.

The high consistency performance reflects the maturity of mainstream cryptocurrency exchanges in price discovery mechanisms. The maximum price difference among the three platforms typically does not exceed 0.002%, with these minor differences primarily stemming from subtle variations in market microstructure rather than systematic price deviations. As the market leader, Binance's prices often serve as reference benchmarks for other platforms, explaining its superior consistency scores.

4.3.2 Data Freshness Performance Differences

Data freshness emerged as the key factor differentiating service quality among the three exchanges, with platforms showing significant differences in this dimension. Binance achieves 83.53 points with an average delay of 0.878 seconds; Bitget scores 77.95 points with an average delay of 1.066 seconds; MEXC scores only 69.06 points with an average delay of 1.362 seconds. The freshness score gap among the three exchanges reaches 14.47 points, constituting the primary cause of overall DQI differences, as illustrated in Figure 3.

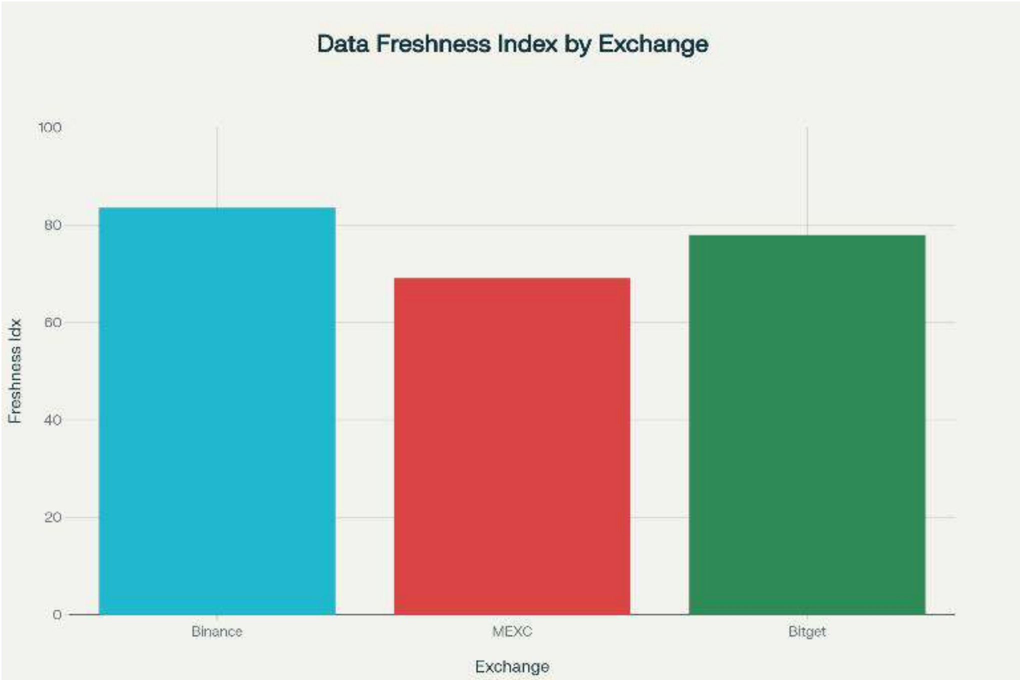

Figure 3: Data Freshness Index by Exchange

Delay time distribution analysis reveals that Binance's delay times are relatively concentrated, with 90% of samples experiencing delays within 1.2 seconds; Bitget's delay distribution is more uniform, ranging between 0.5-1.5 seconds; MEXC's delay times show greater volatility, with some samples exceeding 2 seconds. These delay differences are primarily attributed to variations in technical architecture, server performance, and network infrastructure investment levels across different exchanges.

According to the evaluation framework's segmented penalty mechanism, MEXC's average delay of 1.362 seconds, compared to the ideal 0.2-second standard, resulted in significant score reductions. Such delay differences may create substantial competitive disadvantages in high-frequency trading environments, affecting the capture of arbitrage opportunities and timely trade execution.

4.3.3 Time Series Stability Analysis

Throughout the 8-hour observation period, all three exchanges demonstrated good temporal stability in DQI performance, without apparent cyclical fluctuations or sudden changes. Binance's DQI scores fluctuated slightly within the 92.1-92.8 range with a standard deviation of only 0.23 points; Bitget's score fluctuation range was 88.2-89.1 points with a standard deviation of 0.28 points; MEXC's fluctuation range was 85.0-85.8 points with a standard deviation of 0.24 points.

Comparing performance across Asian and European session divisions, the differences among exchanges between different sessions were minimal. Binance's DQI difference between the two sessions was only 0.6 points, while Bitget and MEXC showed session differences of 0.6 and 0.7 points respectively, all within statistical error ranges. This cross-session consistency indicates that mainstream exchanges' technical infrastructure can maintain stable service levels under different trading activity levels and network load conditions.

Time series analysis also found that the relative rankings among exchanges remained stable throughout the observation period, without rank alternations. This ranking stability reflects the persistence of technical capability gaps among different platforms and validates the reliability and consistency of the evaluation framework.

V. DISCUSSION & CONCLUSIONS

This study successfully established a comprehensive data quality assessment framework for cryptocurrency exchanges, revealing significant service quality disparities that directly impact information fairness. The 7.14-point DQI gap among mainstream platforms demonstrates that even established exchanges exhibit meaningful differences in data service quality, with data freshness emerging as the critical differentiating factor.

The research makes substantial contributions to sustainable digital financial infrastructure development by directly supporting UN Sustainable Development Goals 9.1 and 10.5. The standardized evaluation framework provides quantitative tools for identifying infrastructure weaknesses and promoting more equitable access to high-quality financial data. By establishing transparent assessment standards, this study advances information fairness and reduces systemic inequalities arising from technological disparities among platforms.

Practical implications extend across multiple stakeholder groups. Exchange operators can utilize the framework to identify specific improvement areas and guide infrastructure investments. Regulatory agencies gain objective tools for establishing industry-wide standards and monitoring mechanisms. Market participants benefit from transparent service quality comparisons that inform platform selection decisions. The public availability of our complete dataset and implementation codefurthers collaborative improvement of digital financial ecosystems.

Future research should expand temporal and platform scope to enhance generalizability, develop real-time monitoring systems for continuous quality assessment, and investigate framework applicability to broader financial technology sectors including decentralized exchanges and traditional institutions adopting digital infrastructure.

The correlation between technical investment and service quality underscores the importance of continued infrastructure development in fostering sustainable financial innovation. As cryptocurrency markets mature and integrate with traditional financial systems, ensuring equitable access to high-quality data becomes increasingly critical for maintaining market integrity and promoting inclusive economic growth.

This research establishes that standardized data quality assessment is essential for building sustainable digital financial ecosystems. By providing objective, reproducible evaluation methods, the study contributes to creating more equitable, transparent, and efficient financial markets that serve all participants while supporting broader sustainable development objectives. The path forward requires continued collaboration among researchers, industry practitioners, and regulatory bodies to refine methodologies and promote best practices that advance both technological innovation and sustainability goals.

VI. ETHICAL APPROVAL

This study does not involve human participants or animal subjects. All data was collected through publicly accessible API interfaces provided by cryptocurrency exchanges (Binance, MEXC, and Bitget) for research purposes. The data collection process accessed only publicly available market information without requiring user authentication or accessing personal data. Therefore, ethical approval was not required for this study.

VII. DATA AVAILABILITY STATEMENT

Underlying data:
GitHub: Binance and MEXC and Bitget price data analytic.
https://github.com/kuan35/Binance_and_MEXC_and_Bitget_price_data_analytic

This project contains the following underlying data:

- Bitcoin (BTC/USDT) price data collected from Binance, MEXC, and Bitget exchanges
- Timestamp records for API requests and exchange responses
- Data quality metrics (consistency, freshness, and availability scores)
- Complete dataset of 463 samples collected over 8 hours on September 6, 2025

Extended data:
This project also includes:

- Python scripts for automated data collection via exchange APIs
- Data processing and analysis code
- Documentation for data collection methodology

VIII. GRANT/FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. The Review of Financial Studies, 1(1), 3-40.
[2] Alexander, A. (2025). Latency arbitrage in cryptocurrency markets: Analyzing execution speeds & liquidity dynamics [Working paper]. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5143158
[3] Arner, D. W., Barberis, J., & Buckley, R. P. (2015). The evolution of Fintech: A new post-crisis paradigm? Georgetown Journal of International Law, 47, 1271-1319.
[4] BlockBeats. (2025, September 10). 加密市場全面復甦，總市值回升突破 4 兆美元. 鉅亨網. https://m.cnyes.com/news/id/6150456
[5] Brandvold, M., Molnár, P., Vagstad, K., & Valstad, O. C. A. (2015). Price discovery on Bitcoin exchanges. Journal of International Financial Markets, Institutions and Money, 36, 18-35.
[6] Corbet, S., Lucey, B., Urquhart, A., &Yarovaya, L. (2019). Cryptocurrencies as a financial asset: A systematic analysis. International Review of Financial Analysis, 62, 182-199.
[7] Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. The Journal of Finance, 46(4), 1325-1359.
[8] Hafner, C. M. (2018). Testing for bubbles in cryptocurrencies with time-varying volatility. Journal of Financial Econometrics, 18(2), 233-249.
[9] Kyle, A. S. (1985). Continuous auctions and insider trading. Econometrica, 53(6), 1315-1335.
[10] Ozili, P. K. (2018). Impact of digital finance on financial inclusion and stability. Borsa Istanbul Review, 18(4), 329-340.
[11] Schwenkler, G., Shah, A., & Yang, D. (2025). Aggregate confusion in crypto market data [Working paper]. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4673905
[12] Tsao, J. K. (2025). Binance and MEXC and Bitget price data analytic [Computer software]. GitHub. https://github.com/kuan35/Binance_and_MEXC_and_Bitget_price_data_analytic
[13] United Nations General Assembly. (2015). Transforming our world: The 2030 agenda for sustainable development.
[14] Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. Journal of Management Information Systems, 12(4), 5-33.
[15] World Bank. (2022). Financial inclusion, digital payments, and resilience: How countries can better prepare for the next crisis. World Bank Group.