International Journal of Emerging Trends in Computer Science and Information Technology
ISSN: 3050-9246 | https://doi.org/10.63282/3050-9246.IJETCSIT-V7I1P109
Eureka Vision Publication | Volume 7, Issue 1, 63-72, 2026

Original Article

Data Normalization and Schema Evolution for Ticker Symbol Insights

Shreyansh Sharma
Senior Software Engineer Texas, USA.

Received On: 04/12/2025 Revised On: 05/01/2026 Accepted On: 13/01/2026 Published On: 26/01/2026

Abstract - In this paper, the author will analyze the manner in which the data normalization and schema evolution enhance the management of ticker symbols in financial systems. It points to the constraints of the conventional flat and poorly normalized designs to manage corporate actions and identifier changes and be historically accurate. The study uses normalization to 3NF plus bi-temporal extensions with expand-migrate-contract schema evolution patterns to show how financial databases can ensure superior data integrity, regulations compliance and scalability of the systems. The results offer effective architectural principles of creating robust, scalable, and quality infrastructures of financial reference data.

Keywords - Ticker Symbol Management, Data Normalization, Schema Evolution, Financial Data Architecture, Identifier Mapping, Data Quality, Regulatory Compliance, Financial Databases.

1. Introduction

1.1. Background and Context

The contemporary financial ecosystem creates huge market data flows by the second, amounting to more than 9 million price changes daily in the world exchanges. Nonetheless, the ticker symbol data has certain peculiarities when it comes to management, which is what makes it very different as compared to a typical database application [1, 2]. Mergers, acquisitions, stock splits, reassigning of the ticker symbols, and corporate restructuring are routine processes of financial instruments. The symbols under which a single company has traded in its history may evolve and historical data should be available and accurately maintained to be used to back test or in the course of compliance and audit. As an example, Citigroup, Inc. trades at the current symbol of C but has had several mapping events such as change of ticker symbols, CUSIP changes throughout the trading history of the company [3]. In the 2025 Triennial Central Bank Survey of foreign exchange (FX) and over-the-counter (OTC) interest rate derivatives markets activity, the trading of foreign exchange markets in April of 2025 had reached 9.6 trillion per day, 28 years more than it was in 2022. OTC interest rate derivative trading increased by 59 per cent to $7.9 trillion a day [36].

There are dire repercussions of bad data in the financial systems. Imprecise ticker mapping may cause a trading error, a disvalued portfolio, a violation of regulations and an enormous loss of money. Recent research shows that data quality concerns in financial institutions lead to operational inefficiencies, poor risk assessment and poor decision-making processes. Financial institutions should therefore formulate sound data management models that are accurate, consistent, complete and timely and which are dynamic to the changing nature of the financial markets.

1.2. Research objectives

The main purpose of the study is to find out the applicability of a systematic analysis of how data normalization and schema evolution techniques could be utilized to better manage ticker symbols in financial databases with the aim of enhancing data integrity, historical accuracy, scalability, and regulatory compliance. The paper aims to show how the principles of traditional normalization can be generalized with the time and change of the schema concepts to capture temporal dynamicity of the financial identifiers influenced by corporate activities, restructuring, and regulation modifications.

To be more precise, the study focuses on establishing a systematic architectural design that will allow financial systems to retain stable, precise, and traceable ticker symbol data over time and provide high-frequency updates, sophisticated analysis questions, and historical preservation.

1.3. Research Objectives

The following are the objectives that the paper will seek to attain:

- To explore the basic concepts of data normalization and how they can be applied to the management of ticker symbols data.
- To study schema evolution strategies capable of adapting to dynamic changes in financial markets and keeping the history intact.
- To investigate the identifier mapping systems (CUSIP, ISIN, FIGI) and their presence to normalize financial instrument identification.
- To determine the effect of a correct data normalization and schema evolution on data quality, query performance and regulatory compliance.

1.4. Research Contributions

This study contributes to the area of financial data management in a number of ways:

- Theoretical Contributions: it presents a holistic model that combines the database normalization theory and the need to identify financial instruments and extrapolate traditional normalization concepts to incorporate the temporal factor and corporate behavior peculiar to the ticker symbol management.
- Practical Contributions: The article provides implementation guidelines to be adopted by financial institutions which are interested in updating their market data infrastructure with architectural patterns, schema designs and migration strategies that strike a balance between the benefits of normalization and performance demands.
- Domain-Specific Insights: We obtain domain-specific knowledge of what variables and anti-patterns to design ticker symbols to, based on the study of real-world challenges in managing ticker symbols, which influence the quality of data and system maintainability in financial applications.

1.5. Paper Organization

The rest of this paper is structured in the following manner: Section II is a literature review that will be rich in terms of database normalization theory, schema evolution methods, and financial identifier systems. Section III reports about the research methodology used in the study. Section IV gives an in-depth discussion of normalization methods used on the ticker symbol data. Section V looks into schema evolution schemes of financial databases.

2. Literature Review

2.1. Database Normalization Concepts

Database normalization, A systematic method of data organization in order to reduce redundancy and dependency problems. The idea of normalization was introduced by Edgar F. Codd in the 1970s as the normalization of the relational database model, which introduced a set of normal forms, each with increasingly stricter data organization rules [4]. The normalization process breaks down large tables into smaller related tables joined by foreign key relationships to ensure that every bit of data is presented in only one location.

The first normal form (1NF) presents the basis by making it mandatory that all columns of the table hold atomic and indivisible values and the rows of the table are uniquely identifiable [5]. This eradicates duplicated records and multi-valued fields establishing a uniform format of data storage and retrieval. In the case of ticker symbol data, 1NF is used to provide a single row in the base table per security, and all attributes do not take the form of lists or arrays.

Second Normal (2NF) expands on 1NF by removing partial dependencies, i.e. all non-key attributes must depend on the entire primary key as opposed to only a portion of it [6]. This especially applies to composite keys in which attributes may be reliant on a single key component. In ticker symbol databases using composite keys (symbol + date), 2NF will guarantee that attributes, such as the name of the company, are not based on the date, but rather based on the security identifier.

Third Normal Form (3NF) eliminates the transitive dependencies by having the non-key attributes to rely on the primary key and no other non-key attributes [6]. This also minimizes redundancy and anomaly of updates. In the case of financial information, 3NF is used to separate company data and ticker symbol mappings, so that company data can be stored, and then used by a combination of different ticker symbols in the future.

2.2. Database System Schema Evolution

Schema evolution allows the problem of changing database structures as time moves on whilst still ensuring data is still accessible and the applications still work [7]. With the evolving needs of the business, databases need to be modified by adding new tables or columns, changing the structure of the existing ones, or by depreciating the outdated components. The key problem is to implement these changes without affecting live applications or preventing the loss of the historical data.

Forward Compatibility is used to make sure that systems with newer schemas can read data that has been written with older schemas. This trend is critical in a case where a result can be used by the consumer applications slower than the schema update [8]. New fields can be added by producers and existing consumers blindly ignore the presence of new fields when processing them. The Expand-Migrate-Contract Pattern offers a methodological way of evolving schema that will preserve forward and backward compatibility during the change process.

2.3. Financial Identifier Systems

Securities are represented in a distinct identifier system in financial markets as a way of integrating data between institutions. The presence of several types of identifiers poses a challenge of data normalization and demands the use of advanced mapping functions.

Ticker Symbols are the exchanges that assign short codes that are alpha-numeric symbols to identify publicly traded securities [9]. Although intuitive and commonly known, ticker symbols have severe limitations to data management. Symbols keep on changing with the activities of the corporations, the same symbol can symbolize various securities at various times or on various exchanges and symbols do not carry any geographic or asset class based information. As an example, the ticker symbol C has been used to represent Citigroup on the NYSE but it has been used to represent other entities in the past [10].

CUSIP (Committee on Uniform Securities Identification Procedures), is the issuer of the nine-character alphanumeric code that identifies uniquely financial securities in North America. CUSIP identifiers, created in 1964 to resolve the crunch in the securities settlement paper, are fungible and are the same irrespective of the trading venue [11]. CUSIP

Global Services operates the CUSIP system, and uses it as the National Securities Identification Number (NSIN) of US and Canadian securities. But CUSIP is costly to license and this is a stumbling block to certain organizations.

ISIN (International Securities Identification Number) is a 12-character alpha numeric code that offers international security identification [12]. The initial two characters are used to show the country of issuance and thus, ISIN is of great value in cross-border trading and international portfolio management. Similar to CUSIP, the commercial use of ISIN has to be licensed with fees.

OpenFIGI (also called FIGI, Financial Instrument Global Identifier, FIGI) is a global system of identifying financial instruments in markets and asset classes [13]. FIGI was developed by Bloomberg and is an open standard given to the Object Management Group, which provides a less expensive alternative to CUSIP and ISIN in most applications. The identifier system has various granularity levels, as the identifying may be done at the security level and at the trading venue level.

Identifiers Mapping systems offer essential infrastructure in changing between various identification schemes [14]. These systems keep mappings of history that monitor the changes in symbols, corporate moves and cross identifiers.

2.4. Financial Systems Data Quality

The quality of data is a major issue to financial institutions in which the inaccuracy or omission of data may cause trading errors, breach of compliance, and major financial losses [15]. Studies determine that there are some major dimensions of data quality in regard to ticker symbol management, these are:

Precision so that data is accurate in the sense that they reflect the actual entities and values that they claim to be describing. In case of ticker symbols, accuracy should be that mappings accurately associate symbols with securities, prices should reflect actual market transactions and corporate action data should reflect all the pertinent events [16].

Completeness is the determination of the existence of all the necessary data elements. Unfinished ticker symbol databases need not necessarily have the historical mappings, may not include the most recent corporate actions, or may not have price history gaps that interfere with the analytical functions.

Consistency is the one that guarantees uniformity of data inside and amongst systems and time. The varied reporting standards, the difference in the interpretation of corporate actions, the failure to synchronize systems all result in inconsistent ticker symbol data.

2.5. Financial analytics Data warehousing

Financial Data warehouse schema designs are usually of the star schema or snowflake schema. Star schemas consist of fact tables of measurable events (trades, quotes) and denormalized dimension tables (securities, time, venues) enclosed in star schemas [17]. This design is an optimal query performance structure that reduces joins. Financial analytics data warehousing offers centralized and organized vault which assimilates transactional, market and reference information which involves more than one system inside and outside the enterprise so as to offer consistent and reconcilable perspectives of positions, exposures and profitability throughout the enterprise [18]. Organizations can also manage sophisticated analytical loads needed in risk modeling, stress testing, forecasting and regulatory reporting without straining operational systems by using dimensional modelling using subject oriented fact and dimension tables [19]. The latest financial data warehouses are moving in the direction of using cloud or hybrid-based architecture to enable the elastic scale of compute and storage, lower infrastructure costs, and support increased amounts of historical and real-time data while maintaining a high level of security, governance, and compliance controls.

2.6. Research Gaps and Opportunities

Although the literature has been researched on database normalization, schema evolution, and financial data management separately, there is a paucity of literature that discusses the combined application of the three on ticker symbol management in particular. Key gaps include:

- Absence Domain-Specific Normalization Patterns: The normalization theory does not consider the temporal and versioning requirements that are unique to ticker symbol data.
- Sparse Schema Evolution Recommendations on Financial Systems: The majority of schema evolution work concentrates on general application databases as opposed to the problems of financial reference data that have tight audit and compliance needs [20].
- Lack of Proper Identifier Mapping Integration: The current literature does not consider identifier mapping as an independent issue but instead it does not combine identifier mapping with database design and normalization methods.
- Performance-Quality Trade-offs: There is scant empirical advice available on how to trade off the benefits of normalization with the performance requirements of queries in high-frequency financial systems [21].

The paper will fill these gaps by building a combined framework that integrates the concept of normalization, schema evolution models, and the management of financial identifiers to formulate strong ticker symbol data architectures.

3. Research Methodology

3.1. Research Approach

The research design that has been applied in this study is a mixed-methods study that incorporates theoretical analysis, case study analysis and practical implementation analysis. The approach will combine the positivist philosophy with the interpretive analysis to come up with both generalizable rules and specific guidelines.

3.2. Data Collection Methods

- Literature Review: In-depth examination of scholarly articles, industry standards, technical literature, and best practice manuals with regard to the normalization of databases, schema development, and financial data management. The sources will be IEEE publications, database systems research, financial technology vendors and regulatory guidance.
- Case Study Analysis: The analysis of the actual examples of applications of ticker symbol data management systems of financial institutions, market data vendors, and technology platforms. Case studies involve various situations such as high-frequency trade systems, regulatory reporting systems and portfolio management systems.
- Technical Documentation Review: Review of schema designs, data dictionaries, and system architectures of market data providers Tick Data, Databento as well as financial API services [22]. This review gives real-life examples of the normalization and evolution patterns of production.
- Expert Consultation: The consultation with database architects, financial engineers, and data platform experts to confirm the findings and seek real-life experience about the implementation issues and best practices.

3.3. Analytical Framework

The analytical frameworks used in the research are many:

- Normalization Form Analysis: Analysis of systematic evaluation of ticker symbol data structure against 1NF, 2NF, 3NF and BCNF requirements to identify redundancy and dependencies.
- Data Quality Assessment: Determination of data management methods in accordance with the dimensions of accuracy, completeness, consistency, timeliness, and validity of the data.

3.4. Validation Approach

The validation of research findings is achieved in several ways:

Theoretical Consistency: It is necessary to make sure that the suggested patterns and frameworks are consistent with the already accepted principles of the theory of databases and financial data management.

- **Practical Feasibility:** It must be established that recommendations are feasible within the available technologies and as to within operational limits of financial institutions.
- **Comparative Analysis:** Compare Methods of the proposed approaches with the implemented ones and see how they can be improved and what limitations are possible.

3.5. Research Limitations

Some of the limitations that have been recognized by this study include:

- **Scope Constraints:** The study concentrates on equity ticker symbols mostly with little concentration on derivatives, fixed income and other asset classes which may have different identification requirements.
- **Technology Evolution:** There is a high rate of change in both database technologies and data platforms: Given the changes in capabilities, certain technical recommendations might be updated.
- **Institutional Variability:** The financial institutions vary in their needs, regulatory limits, and system heritage which influence the generalization of recommendations.

3.6. Schema Evolution Design Strategy

In this study, the expand migrate contract schema evolution strategy is taken as the fundamental design principle. The expand stage is the addition of new attributes, tables or structures to the database schema without any changes or deletions of existing elements. This will enable both the new and the old schema to coexist so that the system is not affected in its operation. During the contract stage, the elements of the legacy schema are degraded and only removed when they are proven to have been fully migrated and the process of dependency on the applications have been resolved.

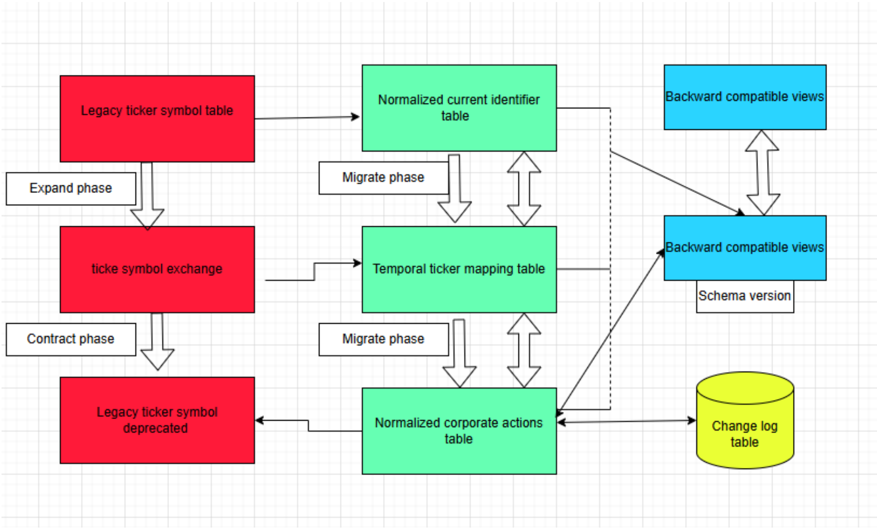


Figure 1. Schema Evolution Design Strategy

4. Techniques of data normalization of Ticker symbol data

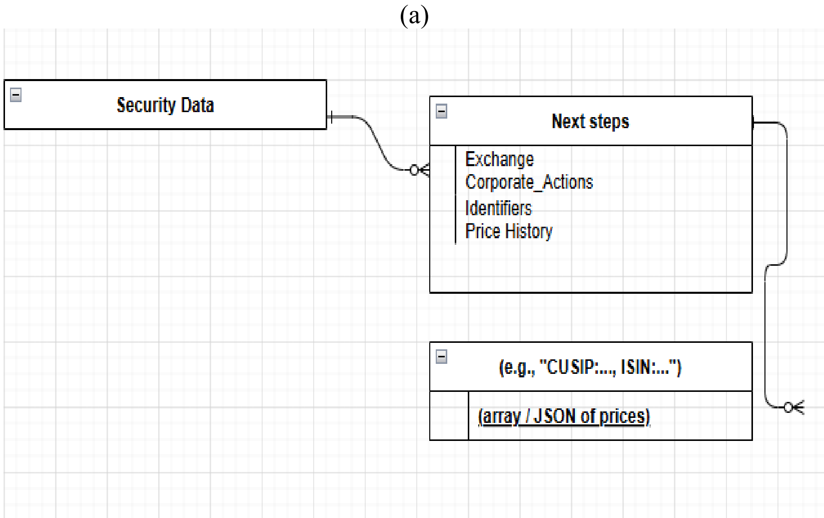
4.1. Problems with Ticker Symbol Data management

The symbol data in ticker symbols has some features that make it hard to use conventional database design methods:

- **Temporal Complexity:** Securities are subjected to corporate activities that form the intricate historical relations. The same company can have several ticker symbols in history, which need to be tracked through time to identify identifiers [23].

```
Security_Data (
  Symbol,
  Company_Name,
  Exchanges [List],
  Identifiers [Map],
  Price_History [Array],
  Corporate_Actions [Array]
)
```

(a)



(b)

4.2. Normalization Trade-offs and Practical Considerations

Although normalization offers advantages of data integrity, it also creates the issue of performance to financial applications [2]. Systems with high frequency trading that query millions of records per second can have delays due to joins between highly normalized structures. Studies indicate that there are a number of strategies to strike a balance between normalization and performance:

- Caching Strategies: Cache popular reference data, in memory, so as not to go through the database repeatedly to join the tables.
- Hybrid Solutions: Operational transaction processing systems (OLTP) should use 3NF, whereas analytical systems (OLAP) should use dimensional modelling with controlled demoralization [24].

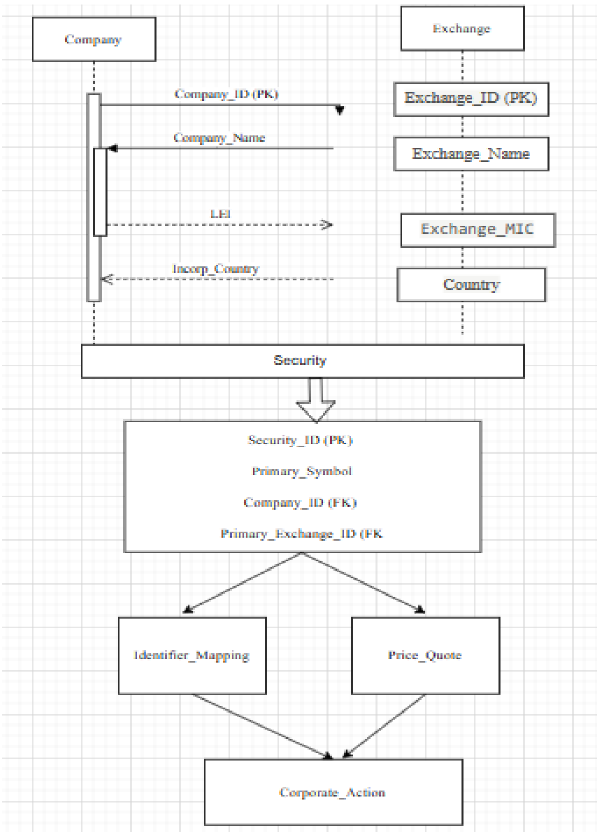


Figure 2. Normalization of Ticker Data (A) Un-Normalized / Poorly Normalized Table (B) Normalized 1NF–3NF Schema

5. Schema Evolution Strategies Financial Database Strategies

5.1. Schema Evolution in Financial Systems Requirements

Uniqueness Financial databases have special schema evolution needs that are fueled by regulatory requirements, market structure changes and technological changes [25]. Key requirements include:

- Zero-Downtime Changes: Since the trading systems are 24/7 across all global markets, there are no schema change maintenance windows.
- Backward Compatibility: When using the same database, several versions of the application can co-exist and both the old and the new code must be supported through a schema [25].
- Prudent Accountable Audit Trial: The documentation of any change made to the schema should be in a way that it can be reversed to facilitate regulatory audit.
- Availability of historical Data: Schema evolution may not impact availability of historical data needed to satisfy compliance and back testing.

5.2. The Expand-Migrate-Contract Pattern

This trend is especially useful with ticker symbol data when the changes in symbols, corporate activities, and mapping updates of identifiers are frequent [26].

- Migrate Phase: The migrate phase will move data and application code and use new structures. This stage will have two writes and the applications will write to an old schema and a new schema.
- Contract Phase: The contract phase eliminates outdated elements after migration has taken place and all the applications operate on the new schema [27].
- Handling Breaking Changes
- Breaking changes are needed to make some schema changes that cannot be fully compatible. Breaking changes can include in the case of ticker symbol data:
- Modifications of important Structure: Modifying the definition of primary key or unique constraints.
- Relationship Modifications: Modifying foreign key relations or referential integrity concepts.

6. Case Study and Implementation Guidelines

6.1. Case Study 1: The High-Frequency Trading System

- Background: The company is a quantitative trading firm that receives millions of updates to the market information every second in equity markets worldwide. The system demands sub-milliseconds of trading decisions and shows the right historical data to support back testing algorithms [28].
- Problem: The first installation employed denormalized flat files, which replicated billions of records of ticker symbol metadata, which consumed too much storage and created synchronization

problems whenever a corporate action took place [29].

- Operational Cache: Denormalized cache In-memory cache: preloads all symbols which have all the metadata and are frequently accessed.
- Historical Archive: Analytical selective-denormalized columnar storage (Parquet)
- Findings: Storage demand reduced by 65 by doing away with redundant data. Optimized denormalization increased the query performance by 40% when the query was being back tested. The time of processing corporate updates dropped to a few minutes due to normalized update mechanisms.

6.2. Case study 2: Regulatory Reporting Platform

- Background: A multinational investment bank needs to adhere to several regulatory frameworks demanding the proper reporting of historical ticker symbols in transaction reporting, position reporting and market surveillance.
- Identifier Resolution Service: Service used to map between all types of identifiers and have a temporal tracking. Systematic tracking increased identifier mapping accuracy to 99.7% as compared to 94% [37]. None of the regulatory breaches on data accuracy in 24 months after implementation.
- Thesis: Regulatory compliance has to be bi-temporally designed regardless of performance implications [30]. Audit trails and schema versioning are impossible to compromise. The benefits of making investments in inclusive identifier mapping infrastructure are realized in a variety of applications.
- Delta Lake Implementation: Goods: Immutable audit log and time travel.

6.3. Case Study 3: Multi-Asset Portfolio Management

- Background: An asset management company is responsible for portfolios in equities, fixed income, derivatives and other alternative investments and it is necessary to have a single identifier management in the various classes of assets.
- Challenge: The incompatible identification schemes were applied to different asset classes. The ticker symbols were used in equities, the CUSIP in bonds, and the OCC in derivatives [31]. There was no consistent perspective on security holdings which made it difficult to manage the risks and report about the performance. API Layer: RESTful API which provides the collective access to the security data including the translation of identifiers.
- Results: Firm-wide perspective of holdings met the first time, allowing to comprehensively aggregate risks. The time spent onboarding new data sources went down to weeks as a result of standardized integration framework. The quality of data was also increased by centralized normalization and validation (78% to 96%).

7. Challenges and Best Practices

7.1. Performance Optimization Challenges

One of the main issues in rigorous normalization of ticker symbol data is its performance impact, since high normalization systems in general may involve multiple joins between large tables, which may introduce quantifiable latency that is intolerable to real-time trading or interactive analytics loads. High-frequency conditions also favor in-memory caching layers on hot reference data and usage of read-optimized denormalised structures and write-optimized normalized master data, such that integrity and speed become balanced rather than being traded off absolutely [32]. The other challenge emerges where several applications and teams need to discuss schema adjustments; a disrespectful penal development course can cause the entire breakdown of downstream work and report code with the slightest modification to an identifier table or symbol history structure.

7.2. Schema Evolution Challenges

The preservation of historical accessibility as the schema change presents its own challenges, especially when they are needed in regulatory scenarios that need to be rebuilt to know an as-was state as of a particular date [32]. Bi-temporal designs (where tables include both valid-time and transaction-time attributes) offer a powerful solution but add to the complexity of the models and the cost of storage, so tend to be used selectively on key reference and mapping tables including symbol history and identifier relationships. Besides structural issues, there has always been a persistent challenge with data quality, because ticker symbol ecosystems are dependent on many external sources which might disagree on identifiers or corporate actions or effective dates [33]. The centralized identifier mapping services and master data management programs have been especially beneficial, as they minimize the mapping logic replication, enhance the data lineage understanding, and offer one area whereby the business rules regarding symbol, CUSIP, ISIN, and FIJI resolution can be controlled and audited.

7.3. Data Quality Challenges

Scalability concerns are experienced in cases when the volume of data and the concurrency requirements increase, particularly where the need to store historical tick data, corporate activities, and cross-asset mappings across the identical ecosystem are necessary [34]. To scale successfully, companies are moving towards cloud-native systems and distributed storage systems with a decoupled compute and storage system so that query engines can be scaled independently and lower-cost archival tiers store colder, less relevant historical data. A clear design of data lifecycle documents like tiring older snapshots into less expensive storage and retaining logically whole views to satisfy regulatory requirements is useful in controlling costs without violating regulatory requirements. Such best practices are the use of role-based access control based on least-privilege, the use of audit trails (preferably append-only) on both data and schema modifications, and data sensitivity classification to ensure the adoption of different retention and masking policies in an organized, documented manner [35]. A coherent combination of these technical, organizational and regulatory practices allows firms to reach a normalized and evolvable architecture of ticker symbols that is performant, credible and compliant to the ongoing change.

8. Conclusion

8.1. Summary of Key Findings

In this study, the extensive usage of data normalization and schema evolution techniques in ticker symbol data management of financial systems has been studied. The paper has a number of significant contributions:

Normalization Framework: We determined that ticker symbol data are well suited to be normalized to Third Normal Form (3NF) reference data, but that due to the temporal nature of the data, bi-temporal extensions are necessary. This normalization eradicates data redundancy, avoids update anomalies, and guarantees data integrity and supports the many complex relationships involved in the financial identifier management.

Schema Evolution Strategies: The expand-migrate-contract pattern offers a powerful approach of schema development of ticker symbols without service failure. This design has backward compatibility, rollback functions and can support the zero-downtime needs of 24/7 trading activities.

Identifier Integration: The management of identifier symbols is an advanced infrastructure that normalizes identifier symbols which are ticker symbols, CUSIP, ISIN and FIJI identifiers. ETL and Data Warehousing: Financial data warehouses have the advantage of having hybrid solutions, both of which are the use of normalized reference data and selectively denormalized analytical structures. Star schemas that have slow changing dimensions are a good trade off of query response and data integrity when an analytical workload is involved.

8.2. Practical Implications

In the case of a financial institution adapting or developing ticker symbol data infrastructure, this study recommends a number of practical solutions:

Normalization is beneficial to data integrity and the performance costs in most cases of reference data. Record any alterations, have roll back processes and impose backward compatibility. The centralized mapping lessens duplication, enhances accuracy and eases the maintenance. Actually query patterns in the profile and optimize certain bottlenecks instead of speculatively losing data integrity. Strict data normalization and controlled schema change of ticker symbol data has immediate implications on the functioning of financial institutions, risk and compliance. Well normalized reference data (security, identifier, corporate actions) minimizes duplication and eradicates update anomalies, enhancing consistency between trading system, risk system and reporting system. Such formal evolution patterns as expandmigratecontract reduce downtime and integration breakages in case a new identifier (such as FIGI) or an attribute is being added, allowing safer and more rapid change cycles. All in all, these practices reduce operational risk, streamline auditing, and establish a more flexible data base on which new products, new regulations, and advanced analytics initiatives can be developed.

8.3. Limitations

This study recognizes that there are a number of limitations that constrained the generalization of results:

- Asset Class Scope: The emphasis on equity ticker symbols does not permit generalization to other asset classes such as derivatives, fixed income, and structured products which use different identifier systems and have temporal characteristics.
- Technology Context: Particular technical recommendations will indicate the present condition of database and data platform technologies. The rapid development in this field demands a regular review of technical practices.
- Organization-Specific Factors: There are individual financial institution combinations of regulatory requirements, legacy system limitations and performance requirements that may require the generalized recommendations be deviated.

8.4. Future Research Directions

This work opens up a number of promising points of future research:

- Multi-Asset Extension: Liberate the normalization model (equity) to include derivatives, fixed income and other assets that have their own identifier schemes and time-varying properties.
- Real-Time Processing: Study methods of normalization of stream processing architectures in which the data is received in real-time and transformations are required with a minimum of latency.

8.5. Concluding Remarks

Data normalization and schema evolution are some of the disciplines which when used correctly in managing the ticker symbols in financial institutions, would greatly improve the quality of data, maintainability of the system, as well as the analytic ability. The study shows how ethical use of these methodologies, tailored to the needs that arise in the context of financial markets, will help organizations to create scalable, secure, and compliant data infrastructure.

The relevance of sound data management frameworks is only going to grow as the financial markets keep becoming automated, increasingly more algorithmic in nature, more regulated and intricate. Companies that spend on appropriate normalization, schema evolution patterns, and identifier management framework are in a better place to adjust to a shifting need faster without compromising the data sanctity on which prudent financial choices are founded.

The combination of the modern data platforms with the native schema evolution support, along with the scale of the cloud computing, is indicative that the trade-offs in the past between the normalization and the performance will not be as restricting. These technologies should be adopted in financial institutions without losing the art of principled database design to develop data infrastructures that are able to sustain the present and future innovation.

References

[1] T.-T. Ho and Y. Huang, “Stock Price Movement Prediction Using Sentiment Analysis and CandleStick Chart Representation,” *Sensors*, vol. 21, no. 23, p. 7957, Nov. 2021, doi: https://doi.org/10.3390/s21237957.

[2] M. Pykälistö, “How database normalization and schema versioning improve system efficiency, data integrity, and flexibility in changing business settings,” *aaltodoc.aalto.fi*, 2024, Available: https://aaltodoc.aalto.fi/items/a7b0ec3a-10ee-4479-8b69-9cbfd2809796

[3] City Group, “Citigroup (C) Laps the Stock Market: Here’s Why,” *Nasdaq.com*, 2025. https://www.nasdaq.com/articles/citigroup-c-laps-stock-market-heres-why-1 (accessed Jan. 12, 2026).

[4] E. Codd, “FURTHER NORMALIZATION OF THE DATA BASE RELATIONAL MODEL,” Aug. 1971. Available: https://forum.thethirdmanifesto.com/wp-content/uploads/asgarosforum/987737/00-efc-further-normalization.pdf

[5] N. Antonio and R. Amato, “Mastering database normalization: A comprehensive exploration of normal forms,” Oct. 2023. Available: https://www.researchgate.net/profile/Nicola-Antonio-Roberto-Amato/publication/374509386_Mastering_database_normalization_A_comprehensive_exploration_of_normal_forms/links/652106cffc5c2a0c3bbe361d/Mastering-database-normalization-A-comprehensive-exploration-of-normal-forms.pdf

[6] Z. Zhang and S. Link, “Synthesizing Third Normal Form Schemata that Minimize Integrity Maintenance and Update Overheads: Parameterizing 3NF by the Numbers of Minimal Keys and Functional Dependencies,” *Proceedings of the ACM on Management of Data*, vol. 3, no. 3, pp. 1–25, Jun. 2025, doi: https://doi.org/10.1145/3725362.

[7] Z. Brahmia, F. Grandi, and B. Oliboni, “A Literature Review on Schema Evolution in Databases,” *Computing Open*, vol. 02, Jan. 2024, doi: https://doi.org/10.1142/s2972370124300012.

[8] J. Edwards, T. Petricek, van, and G. Litt, “Schema Evolution in Interactive Programming Systems,” *The Art Science and Engineering of Programming*, vol. 9, no. 1, Oct. 2024, doi: https://doi.org/10.22152/programming-journal.org/2025/9/2.

[9] S. Frischbier, M. Paic, A. Echler, and C. Roth, “Managing the Complexity of Processing Financial Data at Scale - An Experience Report,” *Springer eBooks*, pp. 14–26, Nov. 2019, doi: https://doi.org/10.1007/978-3-030-34843-4_2.

[10] Yahoo Finance, “Citigroup, Inc. (C) Stock Price, Quote, History & News,” *@YahooFinance*, 2019. https://finance.yahoo.com/quote/C/

[11] Investors, “Committee on Uniform Securities Identification Procedures (CUSIP) | Investor.gov,” *Investor.gov*, 2026. https://www.investor.gov/introduction-investing/investing-basics/glossary/committee-uniform-securities-identification (accessed Jan. 12, 2026).

[12] ISIN, “ISIN Organization: international securities identification numbers organization,” *Isin.org*, 2023. https://www.isin.org/

[13] Open FIGI, “Overview | OpenFIGI,” *Openfigi.com*, 2025. https://www.openfigi.com/about/overview (accessed Jan. 12, 2026).

[14] C. Mazzocca, A. Acar, Selcuk Uluagac, R. Montanari, Paolo Bellavista, and M. Conti, “A Survey on Decentralized Identifiers and Verifiable Credentials,” *IEEE Communications Surveys & Tutorials*, pp. 1–1, Jan. 2025, doi: https://doi.org/10.1109/comst.2025.3543197

[15] Z. B. Yusof, “The Role of High-Quality Data in Risk Assessment: Strategies for Ensuring Accuracy, Completeness, and Timeliness in Financial Predictive Analytics,” *International Journal of Advanced Computational Methodologies and Emerging Technologies*, vol. 15, no. 2, pp. 8–16, 2025, Available: https://owenpress.com/index.php/IJACMET/article/view/2025-02-07

[16] L. Evans, “A Smart Data Ecosystem for the Monitoring of Financial Market Irregularities,” *e-space.mmu.ac.uk*, Sep. 21, 2022. https://e-space.mmu.ac.uk/630409/

[17] M. Golfarelli and S. Rizzi, “From Star Schemas to Big Data: 20 $$+$$ Years of Data Warehouse Research,” *Studies in Big Data*, pp. 93–107, May 2017, doi: https://doi.org/10.1007/978-3-319-61893-7_6.

[18] Vijayan Naveen Edapurath, “Building Scalable Data Warehouses for Financial Analytics in Large Enterprises,” *Philpapers.org*, 2024. https://philpapers.org/rec/NAVBSD-2 (accessed Sep. 17, 2025).

[19] N. J. Rodrigues, “Transforming enterprise finance with data-centric architectures and platform integration,” *World Journal of Advanced Engineering Technology and Sciences*, vol. 15, no. 3, pp. 937–945, Jun. 2025, doi: https://doi.org/10.30574/wjaets.2025.15.3.1001.

[20] H. Gadde, “AI-Driven Schema Evolution and Management in Heterogeneous Databases,” *International Journal of Machine Learning Research in Cybersecurity and Artificial Intelligence*, Oct. 19, 2024. https://www.academia.edu/download/119017466/332_356_ijmlrcai_2019.pdf (accessed Aug. 31, 2025).

[21] V. B. Ramu, “Optimizing Database Performance: Strategies for Efficient Query Execution and Resource Utilization,” *International Journal of Computer Trends and Technology*, vol. 71, no. 7, pp. 15–21, Jul. 2023, doi: https://doi.org/10.14445/22312803/ijctt-v71i7p103.

[22] F. Ekundayo, “International Journal of Engineering Technology Research & Management STRATEGIES FOR MANAGING DATA ENGINEERING TEAMS

TO BUILD SCALABLE, SECURE REST APIS FOR REAL-TIME FINTECH APPLICATIONS," 2024. Accessed: Jun. 11, 2025. [Online]. Available: https://ijetrm.com/issues/files/May-2023-22-1747903717-AUG202314.pdf

[23] B. Hamdikatama, "BEYOND ALGORITHMS: AN INTEGRATED APPROACH TO FAKE NEWS DETECTION USING MACHINE LEARNING TECHNIQUES," *JITK (Jurnal Ilmu Pengetahuan dan Teknologi Komputer)*, vol. 10, no. 3, Feb. 2025, doi: https://doi.org/10.33480/jitk.v10i3.6061.

[24] AWS, "OLAP vs. OLTP - Comparing Data Processing Systems - AWS," *Amazon Web Services, Inc.*, 2025. https://aws.amazon.com/compare/the-difference-between-olap-and-oltp/

[25] R. Sekar, "The need for auto schema evolution in modern data engineering: Challenges and solutions," *World Journal of Advanced Research and Reviews*, vol. 26, no. 1, pp. 909–917, Apr. 2025, doi: https://doi.org/10.30574/wjarr.2025.26.1.1115.

[26] Manolis Katopis, "The Expand and Contract pattern," *Medium*, Oct. 24, 2023. https://medium.com/@ekatopis/the-expand-and-contract-pattern-0479a74f4f2d (accessed Jan. 12, 2026).

[27] T. Wellhausen, "Expand and Contract - A Pattern to Apply Breaking Changes to Persistent Data with Zero Downtime," *Tim-wellhausen.de*, 2018. https://www.tim-wellhausen.de/papers/ExpandAndContract/ExpandAndContract.html (accessed Jan. 12, 2026).

[28] CFI, "High-Frequency Trading (HFT)," *Corporate Finance Institute*, 2025. https://corporatefinanceinstitute.com/resources/equities/high-frequency-trading-hft/

[29] F. Musciotto, J. Piilo, and R. N. Mantegna, "High-frequency trading and networked markets," *Proceedings of the National Academy of Sciences*, vol. 118, no. 26, Jun. 2021, doi: https://doi.org/10.1073/pnas.2015573118.

[30] Ravish Tillu, Muthukrishnan Muthusubramanian, and Vathsala Periyasamy, "Transforming Regulatory Reporting with AI/ML: Strategies for Compliance and Efficiency," *Journal of knowledge learning and science technology*, vol. 2, no. 1, pp. 145–157, Jun. 2023, doi: https://doi.org/10.60087/jklst.vol2.n1.p157.

[31] M. A. Mallouli, Romain Perchet, François Soupé, and R. Leote, "Multi-Asset Portfolios with Active and Passive Funds: A Robust Optimization Framework," *SSRN Electronic Journal*, Jan. 2025, doi: https://doi.org/10.2139/ssrn.5863562.

[32] Ramteke Rakesh, "The Role of Caching in Modern System Design," *Medium*, May 30, 2025. https://medium.com/@ramteke.rakesh/the-role-of-caching-in-modern-system-design-a30132d0c438 (accessed Jan. 12, 2026).

[33] J. M. García Lara, B. García Osma, I. Gazizova, and A. Khalilov, "Demand-driven corporate social responsibility: Symbolic versus substantive change after environmental disasters," *Journal of Corporate Finance*, vol. 94, p. 102816, Sep. 2025, doi: https://doi.org/10.1016/j.jcorpfin.2025.102816.

[34] F. C. Mendes, P. Sarna, Pavel Emelyanov, and C. Dunlop, *Database Performance at Scale*. 2023. doi: https://doi.org/10.1007/978-1-4842-9711-7.

[35] R. D. Edwardson, "Implementing Role-Based Access Control (RBAC) for Least Privilege Security," *Medium*, Mar. 02, 2025. https://medium.com/o-m-n-i-navigating-the-new-cyber-era/implementing-role-based-access-control-rbac-for-least-privilege-security-f382b4083c82

[36] BIS, "Global FX trading hits $9.6 trillion per day in April 2025 and OTC interest rate derivatives surge to $7.9 trillion: Triennial Survey," *Bis.org*, Sep. 30, 2025. https://www.bis.org/press/p250930.htm (accessed Oct. 05, 2025).

[37] C. P. Ratnawat, "AI-POWERED STORY CREATION: TRANSFORMING AGILE DEVELOPMENT IN THE INSURANCE INDUSTRY," *INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND INFORMATION TECHNOLOGY*, vol. 7, no. 2, pp. 2464–2473, Dec. 2024, doi: https://doi.org/10.34218/ijrcait_07_02_183.