

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY SURATHKAL, KARNATAKA



CS251 -DATABASE SYSTEMS Research Paper

Under Professor: Annapa B.

BY:

Name: ANCHAL KUMARI

REG: 2110819 (211CS105)

Section: CSE S1

TITLE OF THE PAPER

Improving Database Quality by Applying
Consistency Aspects of Naming Fields and Table

Author names:

Raissa Maringka

Aulia Khoirunnita

Rodney Maringka

Ema Utami

Journal/Conference name: TEPIAN

Year or publication: 2021

LINK OF THE PAPER:

https://www.researchgate.net/publication/358947286_Improving_Database_Quality_by_Applying_Consistency_Aspects_to_Naming_Fields_and_Tables

ABSTRACT:

A database is a structured collection of data that is organized, stored, and managed to facilitate efficient retrieval, manipulation, and analysis of information. It is designed to store large amounts of data in a structured and organized manner, it allows users or applications to access and retrieve specific pieces of data quickly and accurately.

Ensuring high-quality and consistent data within a database is essential for effective data management and information retrieval. This paper proposes a method for improving database quality by applying consistency aspects to naming fields and tables. Consistency in naming conventions enhances a database schema's clarity, understandability, and maintainability. This study introduces a systematic approach to establishing naming conventions that adhere to recognized standards and best practices. We analyze the impact of consistent naming on database quality through a comparative evaluation of two databases: inconsistent naming and consistent naming. The evaluation considers factors such as readability, comprehensibility, and ease of data integration. The results demonstrate that applying consistency aspects to naming fields and tables significantly enhances database quality, leading to improved data management and more effective information retrieval. This research contributes to the field of database design by highlighting the importance of naming conventions and providing a practical framework for ensuring consistency in database schemas.

Problems addressed in the paper:

Using incorrect names can disrupt data integration and tracking processes. Inconsistent naming of fields and tables makes it challenging to establish relationships and dependencies within the database, hindering the integration of data from various sources and performing complex queries. Without clear and consistent naming, maintenance tasks such as debugging, troubleshooting, and modifying schemas become difficult and error-prone.

Inconsistent naming practices in a database not only frustrate users but also have tangible implications for database decision-making and operational efficiency within an organization. To address these issues, this paper proposes the adoption of standardized naming conventions that employ common nouns and words, promoting efficiency, understanding, and practicality in database development. The goal is to mitigate the problems caused by inconsistent practices through the establishment and enforcement of naming standards. This study examines the impact of employing a common naming registry, including factors such as readability, comprehensibility, ease of maintenance, and information integration. The objective is to provide practical, evidence-based recommendations that foster consensus on common names and enhance overall data quality.

Solutions suggested/implemented:

This article proposes and implements solutions for naming inconsistencies in the database to improve the overall quality and usability of the database. The solution is based on the use of consistency for nouns and words. Some of the solutions mentioned or used in this article are:

1) Create standard naming conventions: This article recommends developing and using standard naming conventions that conform to industry standards and best practices. These rules allow for common and understandable database structures by providing a mechanism for naming fields and regular expressions.

2) Document registration meetings: This document highlights the importance of document recording and makes it easy for all parties involved in document creation and management. Clear data ensures consistency of the database process, making recommendations easier to understand and implement.

3) Education and Training: This document provides education and training to data creators, administrators, and users on the importance of similar applications. By raising awareness and teaching about the benefits of naming conventions, organizations can encourage the practice of naming conventions and develop Good knowledge.

4) Functional and user-friendly tools: This article examines the use of mechanical and user-friendly tools that can check the consistency of information records. This tool scans data schemas, identifies inconsistencies, and offers suggestions that match naming conventions. They also help simplify the process of managing names.

5) Analysis and collaboration: This article promotes the integration of analysis and collaboration in database design and development. By engaging multiple stakeholders to review nominations and provide feedback, organizations can ensure consistent practices and address inconsistencies or trends.

6) Continuous Improvement: This article highlights the need for continuous evaluation and improvement of custom lists. The conference name should be updated and adjusted accordingly as new requirements and changes occur in the literature. Periodic reviews and feedback help identify areas for improvement and maintain

consistency in recording practices. Using these solutions, this article aims to improve the quality of your data by promoting uniformity in registration and language. Proposed guidelines improve literacy, access pressure, and data management schemes by providing guidelines, tools, and shared processes to ensure common practices are followed.

Future work proposed:

This article addresses the problem of improving data quality by combining common features in field lists and tables, but there are many areas where further research and study may be needed to expand solutions. Future research will be conducted as follows, for example, Evaluation of the naming pattern: Further research is ongoing to determine the impact and effectiveness of various standard terminology of quality materials. Comparative studies measuring the readability, clarity, and user satisfaction of various brand recommendations can be used to determine which works best in this process. is created by automation tools and validators specified in Analysis, Automation, and Development Tools. Advanced systems may be developed in the future to manage domain names, provide real-time feedback, and interconnect data management systems.

Related problems which are not addressed:

While this article focuses on improving the quality of information using similarity in names and words, there are many issues that have not been directly addressed before but are worth considering in future research. These related topics include:

- 1) Data Semantics and Content naming: Name fields and words based on semantic and contextual relevance. This article focuses on consistency, but future work may explore how concepts and concepts can be integrated into recording rules to improve understanding and data interpretation.
- 2) Multilingual and Multicultural Databases: These deals with naming conventions in multilingual and multicultural environments. Organizations operating in different regions may have difficulties adapting to different languages and cultures. Future research may explore strategies for maintaining name consistency in this environment.
- 3) Legacy Database Systems: Addresses the challenge of maintaining name consistency on legacy systems. Many organizations have conflicting names for their existing records. Future studies may explore strategies to migrate and enable legacy data using similar applications without disrupting existing processes and operations.
- 4) Domain-Specific Naming: Consider domain-specific naming requirements for a particular domain or business.
- 5) Impact on performance and scalability: Investigate the impact of consistent naming on data performance and scalability. Although this article focuses on improving data quality, it is important to understand how name quality affects query performance, indexing, and overall process scalability. Future research may explore optimizations and trade-offs to reduce performance impact. Answering these pertinent

questions will help you better understand the issues and solutions related to document submission meetings. Future research may explore these areas and provide insights and recommendations for better, accessible, and effective information.

Issues if any, in the proposed solution:

While the proposed solutions for improving database quality by applying consistency aspects to naming fields and tables are beneficial, there are some potential issues that should be considered:

Complexity and Flexibility Trade-off: Implementing standardized naming conventions may introduce a level of complexity and rigidity to the database design process. Stricter naming rules and conventions might limit the flexibility to accommodate unique or evolving requirements. Balancing consistency with the need for flexibility and adaptability is essential to avoid overly restrictive naming conventions.

Compatibility with Existing Databases: Introducing consistent naming practices to an existing database with a large-scale and diverse schema can be challenging. Adapting existing databases to conform to new naming conventions may require significant effort and potentially disrupt ongoing operations. Careful planning and migration strategies are necessary to ensure a smooth transition without compromising data integrity and system stability.

Organizational Adoption and Resistance: Encouraging and ensuring the widespread adoption of consistent naming practices across an organization can be met with resistance. Stakeholders may have different perspectives, preferences, or resistance to change, making it challenging to establish and enforce naming conventions uniformly. Addressing organizational dynamics, fostering collaboration, and providing clear communication about the benefits of consistency are crucial to mitigate resistance.

Maintenance and Evolution: Consistent naming conventions require ongoing maintenance and monitoring to ensure adherence and address evolving needs. As databases evolve, new fields and tables may be added, and modifications may be required, potentially impacting the consistency of the naming conventions. Organizations need to establish processes and guidelines for maintaining and updating the naming conventions as the database evolves.

Cross-Domain Compatibility: The proposed solutions might not adequately address the compatibility of naming conventions across different domains or industries. Each industry may have unique requirements or conventions that need to be considered. Ensuring cross-domain compatibility and flexibility in naming conventions can be a challenge that requires careful consideration and domain-specific customization.

User Acceptance and Usability: While consistent naming practices improve database quality, the usability and acceptance by end-users should not be overlooked. The proposed naming conventions should strike a balance between consistency and user-friendliness, ensuring that users can easily understand and navigate the database without undue confusion or complexity.

Considering these issues and challenges in the implementation of consistent naming practices can help organizations develop robust strategies and solutions that address the specific needs and circumstances of their databases and stakeholders.

MY Solution (thought of / implemented fully or partially)

Analyze Existing Naming Conventions: Start by analyzing the current naming conventions used in the database. Identify any inconsistencies, ambiguities, or patterns that may be causing issues.

Define a Standard Naming Convention: Establish a standard naming convention that aligns with industry best practices and the specific needs of the database. This convention should provide clear guidelines for naming fields and tables, taking into account factors such as data type, purpose, and context.

Automated Data Profiling: Implement automated data profiling techniques to analyze the existing data and identify potential naming inconsistencies. This could involve scanning the database for naming patterns that deviate from the established convention and flagging them for review and correction.

Data Cleansing and Standardization: Develop data cleansing processes to standardize existing field and table names according to the defined convention. This may require renaming existing elements, updating associated documentation, and ensuring data integrity throughout the process.

Continuous Monitoring and Improvement: Establish a process for continuous monitoring and improvement of naming practices. Regularly review and assess the effectiveness of the naming convention, gather feedback from users, and make necessary adjustments or updates to address evolving requirements or emerging challenges.

By implementing this solution, organizations can achieve improved database quality through consistent naming practices. It promotes better data understanding, enhances collaboration, simplifies

maintenance tasks, and improves overall database usability and efficiency.

CONCLUSION:

In conclusion, the paper discusses the importance of improving database quality by applying consistency aspects to naming fields and tables. The proposed solution involves establishing clear naming conventions, documenting them, providing training and education, implementing automated naming validation tools, encouraging peer reviews, and continuously refining the conventions.

While the solution provides a solid foundation for achieving consistent naming practices, there are challenges to consider. These include striking a balance between flexibility and consistency, addressing resistance to change within organizations, ensuring compatibility with existing databases, and maintaining the solution's effectiveness over time.

To further enhance the proposed solution, future work is suggested. This includes evaluating the impact of naming conventions on data analysis tasks, exploring automation and integration with data governance, addressing domain-specific naming requirements, and incorporating user-centric design principles.

Additionally, related problems such as data modeling best practices and data quality management should be considered in conjunction with consistent naming practices.

By addressing these challenges and considering the future work proposed, organizations can strive for improved database quality, usability, and efficiency. Consistent naming practices contribute to enhanced data management, data analysis, and decision-making processes, ultimately leading to better outcomes and insights derived from the database.

REFERENCES:

<https://www.neliti.com/publications/344834/improving-database-quality-by-applying-consistency-aspects-to-naming-fields-and#id-section-content>

<https://www.tibco.com/reference-center/what-is-data-quality>

<https://ieeexplore.ieee.org/document/1521138>

http://www.datamartist.com/resources/datamartist-doc-files/V1_3_Documentation/DM-Automating-Data-Profiling-Doc.html