DOI: 10.2195/lj _proc_lehmann_en_202510_01

Towards Improved Data Quality Management Tools in Logistics

Lukas Lehmann[1], Timon Sebastian Klann[2], Daniel Tebernum[2], Alice Kirchheim[1,3]

Fraunhofer Institute for Material Flow and Logistics IML[1]
Fraunhofer Institute for Software and Systems Engineering ISST[2]
Chair of Material Handling and Warehousing, Department of Mechanical Engineering, TU Dortmund University, Dortmund[3]

In today's logistics environment, high-quality data is essential for ensuring efficient processes and sustaining competitiveness. However, missing, erroneous, or duplicate entries in master data often lead to significant business consequences, such as inefficient supply chains, increased operating costs, and poor decision-making. Existing data screening, cleaning and scoring (DSCS) tools for detecting data errors and thus measuring data quality are often cumbersome to use and are not tailored to the specific needs of logistical master data. In this paper, we present design knowledge to guide the development of DSCS tools. We gathered requirements through dedicated workshops and distilled them into a set of actionable design features. To evaluate our design features, we implemented them in a software prototype, which was tested in a usability and multi-case study. Our contribution in form of design features equips logistics practitioners with concrete guidance for creating and implementing effective DSCS tools in their organizations.

[Keywords: design features, logistics master data quality, data quality tool, design science research]

1 Introduction

Good data quality (DQ) is crucial for the success of a company [1] and is increasingly becoming a competitive factor [2]. The manifestations and requirements are multifaceted. Production planning in an ERP system would be unsuccessful without bills of materials, routing sheets, or planning parameters. Similarly, sales processes cannot be digitized without supplier master data, including addresses or payment terms [3, 4]. Accordingly, master data forms the basis for all knowledge in a company and is ideally of a very high quality [5, 6]. However, data quality can decrease rapidly due to missing, erroneous, or duplicate values [1] leading directly to costly consequences and significantly impairing the competitiveness of companies [7, 8, 9]. Faulty or insufficient data can not only damage the operational process and lead to wrong business decisions, but also affect the potential of data-driven optimization and continuous improvement of own processes. A significant factor contributing to this problem is the increasing complexity of IT and data landscapes. The diversity of data storage and management systems complicates interoperability and fosters inconsistencies and errors [10]. It is reported that data is often faulty and requires extensive preprocessing before it can be reused in industrial projects [11]. Particularly small and medium-sized enterprises often do not have the necessary resources or know-how to objectively assess the quality of their data [12]. They often lack the necessary expertise to handle large volumes of data, as well as the knowledge of specific types of errors and their targeted correction. Furthermore, each dataset has its own specific characteristics, meaning that there is no universal analysis of the data; instead, an individual assessment is always required. For this reason, the general focus in the literature has primarily been on the "consulting" scenario, where companies utilize in-house experts or external consultants to clean data sets [13]. As a consequence, the problem in companies with insufficient data quality is compounded in the form of inefficient processes, increased susceptibility to errors and lack of transparency.

Software tools can be used to improve data quality. However, many companies still rely on Excel or other manual access solutions to validate their master data quality [5]. In order to make specialized software tools more successful in companies, they must be designed in such a way that they meet the requirements for achieving their goal. The literature has already produced a collection of design principles that can be used as a guide when developing data quality tools [14]. However, the practical implementation of these design principles often remains uncertain, as they are very abstract [15]. Therefore, we develop more concrete, actionable design features (DFs) and answer the following research question:

DOI: 10.2195/lj _proc_lehmann_en_202510_01

How should domain-specific data analysis software tools be designed to effectively support logistics practitioners in screening, cleaning, and scoring logistics master data?

To answer this question, we conduct three workshops with logistics experts to gather requirements and abstract the design knowledge for a more generalizable application into design features. The design features are used to implement a prototype that supports logistics experts in analyzing, cleaning and scoring tabular data using easy to understand screening workflows and an intuitive interface. We check the intuitiveness of the tool by means of an usability evaluation with participants. Likewise, we check the generalizability of our prototype by applying it to several logistic cases. The prototype is employed as a test vehicle to evaluate the applicability and robustness of our design features.

The remainder of this paper is organized as follows. Section 2 provides an overview of pertinent related work that informs this study. Section 3 outlines the research design. Section 4.1 presents the results in terms of design features. Section 4.2 details the evaluation methodology. Section 5 discusses the implications of the findings and concludes the paper by summarizing its key contributions and suggesting directions for future research.

2 RELATED WORK

The importance of design knowledge for DQ tools has been widely discussed in the scientific field. However, there is still a need for specific design knowledge of how to build tools that provides a quick and simple analysis of data quality along with easy table management and structured issue tacking, particularly for use in the field of logistics. To our knowledge, this type of artifact has not yet been systematically collected and abstracted.

In an action design research project [16], Altendeitering and Guggenberger design and assess a data-cleaning tool on real master data. The artifact instantiates a "DQ mining pipeline" (ingest, preprocess, detect, export) and integrates complementary techniques for outliers, patterns, association rules, and approximate functional dependencies, refined via focus groups and in-situ evaluations. The core contribution is prescriptive design knowledge for DQ tools: combine multiple algorithms, enable powerful filtering/search over large result sets, guide users with a process-based UI, provide algorithmic explanations favoring interpretability, adopt a service-oriented architecture with standard interfaces, and support (or prepare for) real-time monitoring—principles that inform extensible, human-centered DQ tool design.

In addition to these insights, the authors of [17] conduct a design science research study, combining a literature/tool review and expert interviews to formulate meta-requirements (integrating DQ into data sharing; increasing the effectiveness of shared data), and evaluate an instantiated artifact with practitioners at a manufacturing firm. Based on a data quality app as an IDS-compliant data app tightly coupled to the IDS dataspace connector, the authors posit nine prescriptive design principles (DPs) that constitute derived design knowledge for DQ tools in ecosystems.

Culminating and generalizing this line of research, Altendeitering, Guggenberger and Möller [14] contribute to the field of design knowledge for DQ tools by developing empirically grounded, prescriptive design principles for DQ tools in modern, decentralized data landscapes. Using a design science research program spanning three industrial cases, the authors identify four persistent gaps - automation, integrability, standards, and usability - and formulate 15 design principles that are deliberately abstract, technology-agnostic, and context-independent, synthesized via cross-case abstraction to enable general-purpose applicability across domains, architectures, and stages of the data lifecycle.

The existing literature establishes a solid foundation of prescriptive design knowledge for data-quality tools but stops short of systematically detailing and tailoring these insights for logistics applications. This gap signals that, although we understand what good DQ tools should do in general, we lack concrete, domain-specific artifacts and guidelines for logistics practitioners. Therefore, the body of literature should be extended by deriving and validating concrete, compact design knowledge in the form of design features for DSCS tools. This would bridge the gap between abstract principles and practical implementations.

3 RESEARCH DESIGN

This section describes our approach to addressing the research question and challenges. We base our approach on the Design Science Research (DSR) methodology, which deals with the iterative creation and evaluation of man-made artifacts [18, 19, 20]. Given its practical relevance, Design Science Research (DSR) fosters collaboration between industry and academia [21]. DSR yields artifacts at varying levels of abstraction and detail [22]: at the most concrete end lie software instantiations, followed by e. g. methods, frameworks, and architectures of increasing generality [20]. At the highest level, design principles embody abstract design knowledge in a structured form that makes them reusable [15]. But they are often difficult to operationalize due to their generality [15]. Design features occupy an intermediate position, complementing principles with detailed guidance on constructing other types of artifact like software [23, 24, 25]. Since both design principles and design features are themselves artifacts, they are commonly referred to as meta-artifacts [26].

Within our DSR project, our primary artifact comprises a comprehensive set of design features for DSCS tools in the logistics domain, accompanied by a prototypical instantiation to facilitate their evaluation. Our research process is based on the work of Vaishnavi and Kuechler which consists of five individual steps: *problem awareness*, *suggestion*, *development*, *evaluation*, and *conclusion* [20]. We will now describe each of the five steps in terms of our study.

Phase 1: Problem Awareness. Prior to conducting a DSR project, it is essential to define the underlying problem. This process can involve empirical insights from industry or theoretical perspectives drawn from the scholarly literature. Our discussion of the *problem awareness* is done in Section 1. In short, the challenge lies in data quality in logistics and the lack of suitable software tools for analysis. Practical design guidelines are lacking, as the theoretical principles set out in literature are often too abstract to provide domain-specific guidance for real-world applications.

Phase 2: Suggestion. To enhance the development of successful DSCS tools and, by extension, data quality tools more broadly, we propose the development of specific design features (see Section 2). These features extend the design principles established by Altendeitering et al. [14], providing concrete implementation guidance for practitioners [24, 25]. By complementing the overarching design principles, design features simplify the development process and improve overall tool effectiveness [23].

Phase 3: Development. We follow a supportive ex-ante approach, in which our design features are developed before the actual design process [27]. We first conduct a requirements-elicitation phase to gather concrete requirements for DSCS tools. This is done in collaboration with stakeholders who regularly have to perform DSCS tasks. We identified three stakeholders in our partner networks, who where willing to support us (see Table 1 for characterization). Method wise, we decide on workshops to collect implicit design knowledge in the form of notes on a Miro board. We ask the participants to imagine a software tool that helps in getting the DSCS task done. After the workshop, we review the board and derive design features by reflection and abstraction. Here, we closely follow the work of Braun and Clark to guide our abstraction process [28]. The result is a set of design features, that for better reusability, is coded using the formulation by van Aken [29].

Phase 4: Evaluation. To evaluate the proposed design features, a software prototype is implemented over a period of eight months. The prototype is developed iteratively in weekly cycles. A solution architect guides software developers in translating the design features into actionable development tasks. These tasks are reviewed by a logistics domain expert prior to their implementation.

Table 1: Stakeholder characterization

Role	Description	Experience
Logistic scientists	In the field of applied logistics research with a specialization in packaging logistics	Seven years of project experience with intralogistical issues and specialized in logistical master data, data analysis, and data quality for the past four years
Data scientist	With a background in computer science and data processing	Six years of experience working on software projects in logistics. Research focus on the practical application of artificial intelligence and software-based optimization in packaging and load space optimization, and visualization
Logistic consultant	Intralogistics consulting	Ten years of project experience with intralogistics issues and specialized in packaging software, logistical master data, data quality and data evaluation for the past five years

To validate the software and thus the design features, a two-step evaluation is conducted. **First**, a usability evaluation is carried out [30] with 12 participants from the logistics sector, guided by a script to assess user-friendliness, benefits, and consistency. Their experiences and suggestions are documented in a moderated focus group discussion [31] using a collaborative Miro board. The authors then conduct a SWOT analysis to aggregate the statements from the focus group discussion and define the strengths, weaknesses, opportunities, and threats of the design features implemented in the software [32] [33].

Second, the software's suitability for different logistical master data is examined through a multi-case study involving three data sets from e-commerce, metal processing, and spare parts. A characterization of these use cases is presented in Table 2. An expert review assesses the software's functionalities and applicability. The results of the software analysis are then compared with those of the manual analysis conducted with domain experts, leading to conclusions about the reliability of the results. Drawing on the evaluation results, we critically assess whether the proposed design features have yielded a successful DSCS tool. The results of our evaluation can be found in Section 4.2.

Phase 5: Conclusion. In the final *conclusion* phase of the process model according to Vaishnavi and Kuechler [20], authors must critically examine the results and make

Table 2: Characterization of use cases

Use Case	Data Type	No. of Rows	No. of Columns	Analysis Goal
E-Commerce	Static Master Data	Up to 500,000	Up to 10	Identify data gaps, errors and outliers
Metal Processing	Static and Dynamic Data	Up to 3,000	Up to 30	Analyze data quality and identify errors
Spare Parts Logistics	Static and Dynamic Data	Up to 7,000,000	Up to 15	Identify data gaps, errors and outliers

them accessible to the research community. Our assessment of the results can be found in Section 5. By writing this paper, we share our insights with the research community.

4 RESULTS

In the following we present our original results, come from phases 3 (development) and 4 (evaluation) of the research design.

4.1 DESIGN FEATURES

In the following, we present our design features for developing DSCS tools that answer our research question. We identified a total of 11 final design features that were enriched subsequent to the prototype evaluation. We describe the individual design features orienting towards the template of van Aken [29].

DF1: Basic Algorithms. *To get basic and targeted data analysis insights, provide the DSCS tool with simple and robust algorithms for identifying single-column errors, such as empty values, value range checks or duplicate and similarity detection.*

Basic, column-targeted screenings are foundational to data quality because they detect high-frequency, low-complexity defects that disproportionately degrade downstream analytics. Enabling users to scope algorithms to specific columns and subsets ensures precision, reduces noise, and delivers rapid, actionable feedback at scale that is easy to understand. Incorporating further basic checks - e.g., type checks and forbidden-value detection - complements emptiness, range, duplicate checks, and techniques for similarity search [34] to cover the most prevalent issues in logistics master data, as revealed in the workshop. For example, a targeted scan can flag empty SKU identifiers, detect negative dimensions, identify duplicate material numbers, and reject forbidden status codes across very large tables quickly and efficiently. Regarding similarity detection, simple variants of locality-sensitive hashing methods like “Nilsimsa” can be used.

DF2: ML-driven Analytics. *To efficiently get advanced analytical insights and data interpretation, provide the DSCS tool with ML-driven algorithms.*

ML-driven analytics extend beyond pure syntactic validation to uncover semantic inconsistencies, cross-field dependencies, and context-sensitive plausibility issues that rule-based checks miss. Association analysis (e.g., Apriori [35, 36]) can reveal incompatible attribute combinations or rare co-occurrences that signal data defects, while LLM-based components can act as domain-aware interpreters to evaluate plausibility and consistency. For example, a LLM-based screening can highlight unit mismatches or non-sensical descriptions given domain knowledge.

DF3: Enhanced Outlier Detection. *To identify data issues regarding infrequent, unusual or implausible values, provide the DSCS tool with outlier detection algorithms.*

Outlier detection is essential for maintaining consistency and integrity [37] because anomalies may indicate data entry errors, integration problems, or unusual process conditions. Outlier methods can highlight data points that strongly deviate from normal patterns and thus warrant scrutiny. Utilizing OD in a multi-column (multivariate) detection is particularly important in logistics, where interdependent attributes - such as length, width, height, and weight - must be jointly plausible. Good and robust examples for outlier detection algorithms are Isolation Forests [38], DBSCAN [39] and ECOD [40].

DF4: Domain Abstraction. *To ensure a comprehensive understanding of the screening process for logisticians, provide the DSCS tool with domain-specific abstractions of complex algorithms with simple descriptions.*

Domain abstraction bridges complex algorithmic parameters and practitioner understanding, ensuring that logisticians can confidently configure and interpret analyses. The mental model of the stakeholders must be reflected in the tool. Clear explanations, constrained parameter ranges, and intuitive controls (e.g., sliders) reduce configuration errors and promote consistent, reproducible screening. This usability layer accelerates insight generation by allowing non-technical users to operationalize even advanced methods without deep expertise.

DF5: Efficiency. *To ensure quick processing and screening of large volumes of data, provide the DSCS tool with efficient data pipelines.*

Efficiency is critical in logistics settings where master data can be high-volume and high-dimensional, and timeliness directly affects decision-making. Algorithmic and pipeline-level optimizations enable rapid screening and short feedback cycles. Parameterized algorithms enhance adaptability without sacrificing performance, while asynchronous workflow execution decouples heavy computations from the presentation layer. For example, a workflow engine can orchestrate parallel screenings on partitions of large tables and stream intermediate results to the UI for early validation.

DF6: Modularity. *To ensure flexible extensibility for further domain-specific analysis tasks, provide the DSCS tool with a modular design, allowing for custom algorithms to be added.*

A modular architecture future-proofs the platform by allowing domain-specific checks and specialized algorithms to be added without disrupting core functionality. Standardized interfaces and workflow engines can be utilized in order to enable new screenings to be plugged in, reused, and combined consistently.

DF7: Issue Management. *To ensure effective issue management and resolution tracking, provide the DSCS tool with capabilities for specification, classification and assignment of identified issues.*

Powerful issue management transforms raw detections into actionable, auditable outcomes by classifying, prioritizing, and assigning potential errors. Given that many findings are only potential data issues, manual validation workflows are essential to confirm true errors and minimize false positives. Batch handling of similar issue types improves efficiency, while non-destructive tracking preserves original data and supports governance. Comprehensive logging of screening and resolution steps ensures traceability and compliance readiness. For example, all "forbidden value" findings in a specific column can be grouped into a review queue, manually validated, and then tracked with status, assignee, and resolution notes.

DF8: DQ Dashboard. *To ensure comprehensive understanding of the data quality status, provide the DSCS tool with a graphical representation of the screening results.*

A dedicated dashboard provides comprehension and awareness by summarizing data quality across issue classes and datasets. Visual overviews linked to classified issues (as defined in DF7) allow targeted drill-down, validation, and prioritization. This compresses the time from detection to action and supports stakeholder communication.

DF9: Visualization. *To ensure effective data visualization and quick interpretation of DQ results, provide DSCS tools with color coded data tables.*

Color-coded tables leverage preattentive visual cues to accelerate anomaly recognition and reduce cognitive load during validation. By highlighting cells according to issue types or severity, users can scan large datasets and error tables efficiently and focus on the most critical entries.

DF10: Share & Export. *To share the DQ scoring among stakeholders, provide the DSCS tool with capabilities to store and export the categorized DQ results with descriptions of identified issues.*

Sharing and exporting curated, categorized DQ results enhances collaboration, governance, and accountability across technical and non-technical stakeholders. High-level formats (e.g., Word or PDF) make findings accessible, preserving narrative descriptions, tables, and visuals for decision forums. For example, a PDF report can summarize overall DQ, issue counts by class and dataset, include exemplars, and outline further issues descriptions.

DF11: Upload Assistant. *To ensure that users are aware of the data and possess the capability to utilize, manage and combine different datasets, provide the DSCS tool with an upload assistant comprising easy table management for multiple datasets.*

Data quality issues often originate during data integration; therefore, a guided upload assistant that supports joining, renaming, reordering, and omitting columns is vital, as revealed in the workshop. By enabling users to combine multiple tables into coherent datasets and perform bulk order or cross-table consistency checks, the tool reduces schema drift and linkage errors while enabling DQ Scoring over composed sets of data.

4.2 Evaluation

The evaluation of the design features took place on two levels using a prototypical implementation (screenshots available online [41]): On one hand, the usability, functionality, and consistency of the software were assessed in an evaluation (Usability Evaluation). On the other hand, an evaluation regarding the applicability of the software to different use cases was conducted. It was examined whether the software is suitable for various logistical master data sets, varying stakeholders, and different requirements (Multi Case Study). The results of the evaluations are presented below.

4.2.1 Usability evaluation

The evaluation results of the software's usability was predominantly positive. The operation was perceived as straightforward and understandable, especially the clear design of the user interface. Additionally, the graphical representation of the screening results was rated as helpful. However, weaknesses were identified, including poorly visible buttons (e.g., when setting offsets) and unclear or inaccurate terminology. The lack of help texts and info boxes further compounded the challenges faced by users during navigation. In addition, the participants suggested optimizing error messages and providing specific hints for corrections to improve usability.

With the investigation of functionality, the intuitive representation and utilization of processes such as screening and cleaning emerged as a notable strength. The visual support through pie charts, which facilitate the cleaning process, was particularly praised. However, the upload interface presented certain challenges, as it permitted multiple uploads of identical tables and was perceived as unclear concerning the preview function. The *Accept* button caused confusion in several areas, as its meaning and use were not clear enough. This led to user errors and inconsistent results when working through the given script. Furthermore, the complexity of the software was particularly perceived as a hindrance to inexperienced users.

The evaluation concluded with the SWOT analysis (see Figure 1), which demonstrated a variety of strengths of the software including the efficient processing of large amounts of data, flexible extensibility through modules and workflows, and intuitive operation of basic functions. Furthermore, the efficient implementation of complex workflows was positively evaluated. This results in numerous opportunities, particularly regarding its varied applications in different industries, the possibility and potential of integrating artificial intelligence, and ease of use also for non-technical users. However, the complexity of the software, the risk of overwhelm due to complex functions, and the lack of reliability and reproducibility in AI-supported screenings pose certain issues.

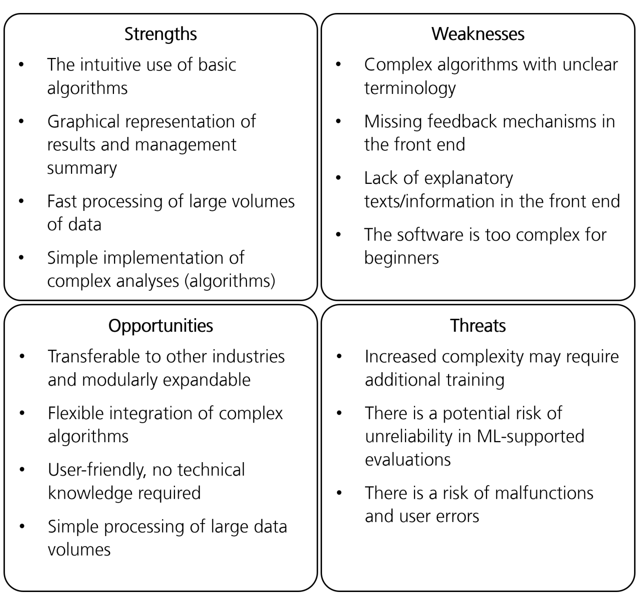

Figure 1: Results of the SWOT Analysis

4.2.2 MULTI CASE STUDY

Static master data in e-commerce: The analysis of a dataset from a garden tools retailer focused on order and article data. The dataset comprised around half a million rows and fewer than ten columns (see Table 2). The main objective was to identify data gaps, errors and outliers in order to ensure the dataset was comprehensive and of high quality. The AI algorithm performed adequately when interpreting the description columns alongside the dimensions (length, width and height) and weight. However, it often misinterpreted data due to insufficient article descriptions, resulting in false positives. Although the software reliably detected outliers, it was not precise enough to be the sole method of assessing data quality. The reliability of data quality assessments was enhanced by manual domain knowledge, and the export function provided a valuable overview of data quality, highlighting incomplete or defective data and identifying error types.

Dynamic transaction data in metal processing: This investigation analysed a dataset from a metal processing company which included static master data and dynamic metrics, amounting to 30 columns in total (see Table 2). The software was user-friendly and effectively executed all the necessary analysis steps, primarily using basic algorithms. The analysis speed was acceptable, enabling new analyses to be initiated during asynchronous execution. An initial upload error caused by line breaks in the source file was resolved by adjusting the upload format. However, there was a notable lack of functionality to save analysis steps as profiles, meaning users had to repeat manual processes, which was time-consuming with extensive datasets. Additionally, the outlier detection methodology was criticised for being unclear, which complicated the interpretation of results. While AI analysis was not necessary for this dataset, the software received a positive overall assessment, with suggestions for improvements to the documentation and efficiency.

Static and dynamic master data in spare parts logistics: This evaluation analysed a dataset from a retailer of spare parts for large electrical appliances, consisting of around seven million rows. This included both static logistical master data (e.g. article designation, dimensions, weight) and dynamic metrics (e.g. consumption quantity) (see Table 2). Initial issues arose during the upload of the xlsx data, which were resolved by converting it to csv format. The software effectively executed all the necessary analytical steps, identifying more errors than manual analysis, particularly with regard to duplicates and measurement errors. While the base algorithms produced results in an acceptable timeframe, outlier detection and AI functions required significantly longer processing times. An analytical error regarding weight data persisted even after multiple executions, indicating a further area of investigation. Overall, although the software successfully detected errors, the long loading times for large datasets highlighted the need for performance optimisation.

In summary, the evaluation shows that the software has robust capabilities for analysing both static and dynamic data and consistently identifies more errors than manual analysis. While it excels in efficiently processing large datasets and offering valuable insights into data quality, it faces challenges related to data interpretation and lacks efficiency features for repeated analyses. Additionally, it experiences performance issues with extensive data. Addressing these limitations could significantly enhance the software's overall utility and user experience. The key identified capabilities and limitations of the software are summarized below:

Capabilities:

- **Comprehensive data analysis:** The software can effectively identify errors and outliers across various datasets, uncovering more issues than traditional manual methods.
- **Efficiency with large datasets:** It can process extensive datasets (e.g. millions of rows), which are typically unmanageable with conventional tools like Excel.
- **User-friendly interface:** Its intuitive design enhances usability, enabling users to perform complex analyses with relative ease.
- **Export functionality:** The software provides export options and management summaries, enabling the streamlined correction of identified data errors in customer systems.

Limitations:

- **Interpretation issues:** The AI algorithm often misinterprets data due to inadequate or unclear article descriptions, leading to false positives.
- **Manual process repetition:** The inability to save analysis profiles means that analysis steps must be repeated manually, which increases the time and effort required, especially for extensive datasets.
- **Outlier detection clarity:** The methodology behind outlier detection is not well-defined, which complicates users' ability to correctly interpret analysis results.
- **Performance challenges:** Long loading times for large datasets suggest areas where the software could be optimized, particularly since it is still in the prototype stage.

5 DISCUSSION & CONCLUSION

In this paper, we aim to answer the question of how to support logistics practitioners in developing and enacting tools for data screening, cleaning, and scoring. To advance the field of data quality management in logistics and achieve this, eleven specific design features for DSCS tools were formulated and operationalized. These features were tailored specifically to logistical master data and were elicited from three workshops with logistics experts (see Section 4.1). We instantiated our design features in a software prototype and validated it through a usability study with twelve practitioners and multi-case evaluations across e-commerce, metal processing, and spare-parts logistics datasets (see Section 4.2).

Our work makes both theoretical and practical contributions to the emerging field of data quality management in logistics. By moving beyond the largely abstract design principles in prior literature, we derive and validate 11 concrete design features tailored to logistical DSCS tools.

From a theoretical standpoint, we add to the body of knowledge by supplementing abstract design principles for data quality tools with a concrete set of eleven design features tailored to the logistical context. By moving from high-level prescriptions to actionable specifications, we bridge the gap between data quality theory and real-world tool development. DF4 (Domain Abstraction) in particular makes a notable theoretical contribution by providing a mid-level construct that bridges the gap between complex algorithmic parameters and the mental models of logistics professionals. This contextual interpretability ensures that even non-technical users can confidently configure, execute, and trust advanced screening methods without having in-depth knowledge of statistics or machine learning. This design feature proved to be very significant in the evaluation of our prototype. DF7 takes a similar approach, which involves the individualization of the classification and assignment of identified data defects. Here, too, the evaluation of our prototype shows that the adaptability of the software to the domain and context is a decisive factor for the success of the tool.

On the practical side, our work delivers a reusable design that software vendors and internal IT teams can immediately apply when developing or extending DSCS tools. The eleven design features summarize best practices in user interface, workflow orchestration, and algorithm selection, and directly address the weaknesses uncovered in our expert workshops. In a usability study with twelve logistics experts and in three different case studies on master data in e-commerce, dynamic transaction data in metal processing, and spare parts logistics, the prototype proved to be intuitive and effective in detecting data defects on a large scale. Consequently, the design features can serve as a blueprint to inform and guide the development of new DSCS tools.

At the same time, our research has several limitations that qualify the evidence for the proposed DFs. First, their evaluation was only implicit via a single software prototype that instantiates all DFs simultaneously; we did not isolate effects via A/B tests or controlled comparisons against alternative designs, which introduces threats to internal and construct validity. Consequently, observed benefits may stem from tool-specific implementation choices and unmeasured interactions among DFs. Second, requirements were elicited in three workshops with only few stakeholders, a small and potentially homogeneous sample that risks selection and researcher-coding bias. Third, the DF catalog is incomplete by construction; we neither claim exhaustiveness nor systematic coverage across logistics subdomains. Finally, AI-supported screenings exhibited reproducibility and performance issues on very large datasets, and robustness across data regimes—as well as security, governance, and integration constraints—was not systematically tested, further constraining generalizability.

The results demonstrate that our prototype systematically uncovers a broader spectrum of data defects than manual methods, accelerates cleansing workflows, and enhances transparency and governance. We therefore see our design features as contributing to the development of useful DSCS tools. Our work contributes both to theory, by extending data-quality frameworks within the logistics domain, and to practice, by providing a blueprint for tool developers.

ACKNOWLEDGMENT

This paper was written as part of the transfer project *DSCS - Data Screening, Cleaning and Scoring*, which was funded by the *Centre of Excellence for Logistics and IT* at the Fraunhofer Institute for Material Flow and Logistics (IML).

REFERENCES

[1] S. Kandel, R. Parikh, A. Paepcke, J. M. Hellerstein, and J. Heer, “Profiler: integrated statistical analysis and visualization for data quality assessment,” in *International Working Conference on Advanced Visual Interfaces, AVI 2012, Capri Island, Naples, Italy, May 22-25, 2012, Proceedings*, G. Tortora, S. Levialdi, and M. Tucci, Eds. ACM, 2012, pp. 547–554.

[2] G. Köksal, İnci Batmaz, and M. C. Testik, “A review of data mining applications for quality improvement in manufacturing industry,” *Expert Systems with Applications*, vol. 38, no. 10, pp. 13 448–13 467, 2011.

[3] A. Haug, J. Stentoft Arlbjørn, F. Zachariassen, and J. Schlichter, “Master data quality barriers: an empirical investigation,” *Industrial Management & Data Systems*, vol. 113, no. 2, pp. 234–249, 2013.

[4] M. Zillmann, “Ohne verlässliche stammdaten geht es nicht,” *Controlling & Management Review*, vol. 61, no. 6, pp. 68–72, 2017.

[5] T. Schäffer and H. Beckmann, “Trendstudie Stammdatenqualität 2013: Erhebung der aktuellen Situation zur Stammdatenqualität in Unternehmen und daraus abgeleitete Trends,” Hochschule Heilbronn, Tech. Rep., 2014.

[6] D. Linstedt and M. Olschimke, *Building a scalable data warehouse with data vault 2.0*. Morgan Kaufmann, 2015.

[7] T. Dasu and T. Johnson, *Exploratory Data Mining and Data Cleaning*. John Wiley, 2003.

[8] T. C. Redman and T. H. Davenport, “Getting serious about data and data science,” https://sloanreview.mit.edu/article/getting-serious-about-data-and-data-science/, 2020, accessed: 25.05.2025.

[9] T. C. Redman, “The impact of poor data quality on the typical enterprise,” *Commun. ACM*, vol. 41, no. 2, pp. 79–82, 1998.

[10] O. A. Adenekan, N. O. Solomon, P. Simpa, and S. C. Obasi, “Enhancing manufacturing productivity: A review of ai-driven supply chain management optimization and erp systems integration,” *International Journal of Management & Entrepreneurship Research*, vol. 6, no. 5, pp. 1607–1624, May 2024.

[11] M. Hameed and F. Naumann, “Data preparation: A survey of commercial tools,” *ACM SIGMOD Record*, vol. 49, no. 3, pp. 18–29, 2020.

[12] S. Singh and J. Singh, “A survey on master data management techniques for business perspective,” *Lecture Notes in Networks and Systems.*, vol. 291, 2022.

[13] P. Wang and Y. He, “Uni-detect: A unified approach to automated error detection in tables,” in *Proceedings of the 2019 International Conference on Management of Data, SIGMOD Conference 2019, Amsterdam, The Netherlands, June 30 - July 5, 2019*. ACM, 2019, pp. 811–828.

[14] M. Altendeitering, T. M. Guggenberger, and F. Möller, “A design theory for data quality tools in data ecosystems: Findings from three industry cases,” *Data & Knowledge Engineering*, vol. 153, p. 102333, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169023X24000570

[15] L. C. Kruse, S. Seidel, and S. Purao, “Making use of design principles,” in *Tackling Society’s Grand Challenges with Design Science - 11th International Conference, DESRIST 2016, St. John’s, NL, Canada, May 23-25, 2016, Proceedings*, ser. Lecture Notes in Computer Science, J. Parsons, T. Tuunanen, J. Venable, B. Donnellan, M. Helfert, and J. Kenneally, Eds., vol. 9661. Springer, 2016, pp. 37–51. [Online]. Available: https://doi.org/10.1007/978-3-319-39294-3_3

[16] M. Altendeitering and T. M. Guggenberger, “Designing data quality tools: Findings from an action design research project at boehringer ingelheim,” in *ECIS 2021 Research Papers*, 2021. [Online]. Available: https://aisel.aisnet.org/ecis2021_rp/95

[17] M. Altendeitering, S. Dübler, and T. M. Guggenberger, “Data quality in data ecosystems: Towards a design theory,” in *AMCIS 2022 Proceedings*, vol. 3, 2022. [Online]. Available: https://aisel.aisnet.org/amcis2022/DataEcoSys/DataEcoSys/3

[18] S. T. March and G. F. Smith, “Design and natural science research on information technology,” *Decis. Support Syst.*, vol. 15, no. 4, pp. 251–266, 1995. [Online]. Available: https://doi.org/10.1016/0167-9236(94)00041-2

[19] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design science in information systems research,” *MIS quarterly*, pp. 75–105, 2004.

[20] V. Vaishnavi and W. Kuechler. (2004) Design science research in information systems. (updated in 2017 and 2019 by Vaishnavi, V. and Stacey, P.); last updated November 24, 2021. [Online]. Available: http://www.desrist.org/design-research-in-information-systems/

[21] P. Runeson, S. Minör, and J. Svenér, “Get the cogs in synch: time horizon aspects of industry-academia collaboration,” in *WISE’14, Proceedings of the 2014 ACM International Workshop on Long-term Industrial Collaboration on Software Engineering, Vasteras, Sweden, September 16, 2014*, R. Dobrin, P. Wallin, A. C. R. Paiva, and M. B. Cohen, Eds. ACM, 2014, pp. 25–28. [Online]. Available: https://doi.org/10.1145/2647648.2647652

[22] H. Wache, F. Möller, T. Schoormann, G. Strobel, and D. Petrik, “Exploring the abstraction levels of design principles: The case of chatbots,” in *WI for Grand Challenges - Grand Challenges for WI, 17. Internationale Tagung Wirtschaftsinformatik (WI 2022), February 21-23, 2022, Nuremberg, Germany*. AISeL, 2022. [Online]. Available: https://aisel.aisnet.org/wi2022/design_science/design_science/6

[23] D. Jones and S. Gregor, “The anatomy of a design theory,” *J. Assoc. Inf. Syst.*, vol. 8, no. 5, p. 19, 2007. [Online]. Available: https://doi.org/10.17705/1jais.00129

[24] H. Meth, B. Müller, and A. Maedche, “Designing a requirement mining system,” *J. Assoc. Inf. Syst.*, vol. 16, no. 9, p. 2, 2015. [Online]. Available: https://doi.org/10.17705/1jais.00408

[25] M. Rhyn and I. Blohm, “Combining collective and artificial intelligence: towards a design theory for decision support in crowdsourcing,” in *25th European Conference on Information Systems, ECIS 2017, Guimarães, Portugal, June 5-10, 2017*, I. Ramos, V. Tuunainen, and H. Krcmar, Eds., 2017. [Online]. Available: http://aisel.aisnet.org/ecis2017_rip/18

[26] J. Iivari, “The IS core - VII: towards information systems as a science of meta-artifacts,” *Commun. Assoc. Inf. Syst.*, vol. 12, p. 37, 2003. [Online]. Available: https://doi.org/10.17705/1cais.01237

[27] F. Möller, T. M. Guggenberger, and B. Otto, “Towards a method for design principle development in information systems,” in *Designing for Digital Transformation. Co-Creating Services with Citizens and Industry: 15th International Conference on Design Science Research in Information Systems and Technology, DESRIST 2020, Kristiansand, Norway, December 2–4, 2020, Proceedings 15*. Springer, 2020, pp. 208–220.

[28] V. Braun and V. Clarke, “Using thematic analysis in psychology,” *Qualitative research in psychology*, vol. 3, no. 2, pp. 77–101, 2006.

[29] J. E. v. Aken, “Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules,” *Journal of management studies*, vol. 41, no. 2, pp. 219–246, 2004.

[30] P. W. Jordan, B. Thomas, I. L. McClelland, and B. Weerdmeester, *Usability evaluation in industry*. CRC Press, 1996.

[31] M. M. Hennink, *Focus group discussions*. Oxford University Press, 2013.

[32] T. Sammut-Bonnici and D. Galea, “Swot analysis,” *Wiley Encyclopedia of management*, pp. 1–8, 2015.

[33] P. Kotler and K. L. Keller, *Marketing Management*. Pearson, 2016.

[34] J. Wang, H. T. Shen, J. Song, and J. Ji, “Hashing for similarity search: A survey,” *arXiv preprint arXiv:1408.2927*, 2014.

[35] R. Agrawal, T. Imieliński, and A. Swami, “Mining association rules between sets of items in large databases,” vol. 22, no. 2, p. 207–216, Jun. 1993. [Online]. Available: https://doi.org/10.1145/170036.170072

[36] R. Agrawal and R. Srikant, “Fast algorithms for mining association rules in large databases,” in *VLDB’94, Proceedings of 20th International Conference on Very Large Data Bases, September 12-15, 1994, Santiago de Chile, Chile*. Morgan Kaufmann, 1994, pp. 487–499.

[37] V. Hodge and J. Austin, “A survey of outlier detection methodologies,” *Artificial intelligence review*, vol. 22, pp. 85–126, 2004.

[38] F. T. Liu, K. M. Ting, and Z.-H. Zhou, “Isolation-based anomaly detection,” *ACM Trans. Knowl. Discov. Data*, vol. 6, no. 1, Mar. 2012.

[39] D. Deng, "Dbscan clustering algorithm based on density," in *2020 7th international forum on electrical engineering and automation (IFEEA)*. IEEE, 2020, pp. 949–953.

[40] Z. Li, Y. Zhao, X. Hu, N. Botta, C. Ionescu, and G. H. Chen, "Ecod: Unsupervised outlier detection using empirical cumulative distribution functions," *IEEE Transactions on Knowledge and Data Engineering*, vol. 35, no. 12, pp. 12 181–12 193, 2022.

[41] D. Tebernum, T. S. Klann, and L. Lehmann, "Screenshots of the dscs software," Jul. 2025. [Online]. Available: https://doi.org/10.5281/zenodo.16088920

Lukas Lehmann, M.Sc., Research Assistant at the Fraunhofer Institute for Material Flow and Logistics IML.
Phone: +49 231 9743 318
E-Mail: lukas.lehmann@iml.fraunhofer.de

Timon Klann, B.Sc., Research Assistant at the Fraunhofer Institute for Software and Systems Engineering ISST.
E-Mail: timon.klann@isst.fraunhofer.de

Daniel Tebernum, Dr.-Ing., Post-Doc at the Fraunhofer Institute for Software and Systems Engineering ISST.
Phone: +49 231 97677 462
E-Mail: daniel.tebernum@isst.fraunhofer.de

Alice Kirchheim, Prof., Institute Director at the Fraunhofer Institute for Material Flow and Logistics IML and Professor at the Chair of Material Handling and Warehousing, TU Dortmund University
Phone: +49 231 9743 600
E-Mail: alice.kirchheim@tu-dortmund.de, alice.kirchheim@iml.fraunhofer.de

Adresses:
Fraunhofer Institute for Material Flow and Logistics IML, Joseph-von-Fraunhofer-Straße 2-4, 44227 Dortmund, Germany.
Fraunhofer Institute for Software and Systems Engineering ISST, Speicherstrasse 6, 44147 Dortmund, Germany.
Chair of Material Handling and Warehousing, TU Dortmund University, Joseph-von-Fraunhofer-Str. 2-4, 44227 Dortmund, Germany.