

MSQ Methodology in AI-driven Analysis of Assets and Liabilities: A Case Study of the ValidoAI System

Slavoljub Milojković^{1,2,*}, Pavle Dakić^{2,3,6}, Tjaša Heričko⁴, Marijana Aleksić⁵ and Ján Lang⁶

¹Faculty of Business Studies and Law, MB University, Teodora Drajzera 27, 11000 Belgrade, Serbia

²ValidoAI, Cara Lazara 15, 34000 Kragujevac, Serbia

³Faculty of Informatics and Computing, Singidunum University, Danijelova 32, 11000 Belgrade, Serbia

⁴Faculty of Electrical Engineering and Computer Science, University of Maribor, Koroška cesta 46, 2000 Maribor, Slovenia

⁵Faculty of Business Studies and Law, MB University, Teodora Drajzera 27, 11000 Belgrade, Serbia

⁶Institute of Informatics, Information Systems and Software Engineering, Faculty of Informatics and Information Technologies, Slovak University of Technology in Bratislava, Ilkovičova 2, 842 16 Bratislava, Slovakia

Abstract

Small and medium-sized enterprises (SMEs) in Serbia often lack access to tools that enable them to interpret basic financial statements independently, making it difficult to assess liquidity and debt levels. This paper presents the ValidoAI system, which applies a Minimum Sufficient Quantity (MSQ) approach to automate balance sheet analysis using real-world SME data. The core component, the AI Ledger module, transforms unstructured accounting records into structured representations, visualizes asset and liability structures, and generates narrative explanations of financial positions. Emphasis is placed on the relationship between short-term liabilities, equity, and current assets, enabling the system to classify financial stability and highlight potential risks. The methodology integrates a Python-based ETL pipeline, automated data anonymization, and GPT-driven narrative generation, ensuring both transparency and accessibility for non-expert users. To evaluate the system, we have created an example testing dataset by extracting anonymized balance sheet and transaction data from SME accounting software in both PDF and Excel formats. The data was standardized, categorized, and structured to reflect typical SME financial records, enabling robust testing of the automated analysis pipeline. This study is focusing on the potential of AI-driven, explainable systems that can support financial decision-making in the SME sector, even in the absence of formal accounting expertise. Results demonstrate that a small set of well-selected indicators can reliably identify liquidity risks and unbalanced capital structures, providing actionable recommendations without manual intervention.

Keywords

Artificial Intelligence, Financial Analysis, Balance Sheet, MSQ Methodology, Small and Medium Enterprises, Automation

1. Introduction

Small and medium-sized firms (SMEs) in Serbia frequently struggle to evaluate their own financial accounts [1]. Primarily this is due to a lack of financial awareness, inadequate digital tools, and a scarcity of competent labor. This is particularly true for small enterprises, which often operate without structured finance departments or access to professional financial advisory services [2, 3].

According to recent reports from the Regional Cooperation Council [3, 4], small and medium-sized enterprises (SMEs) in Southeast Europe continue to lag behind in digital adoption and the use of analytical tools, call centers limiting their ability to make data-informed financial decisions and hampering their growth potential [5, 6].

To address these limitations, the ValidoAI system with its various modules was developed as a lightweight decision-support tool that interprets the balance sheet using a reduced but targeted set

SQAMIA 2025: Workshop on Software Quality, Analysis, Monitoring, Improvement, and Applications, September 10–12, 2025, Maribor, Slovenia

*Corresponding author.

✉ milojkovics@yahoo.com (S. Milojković); pdakic@singidunum.ac.rs (P. Dakić); tjasa.hericko@um.si (T. Heričko); marijana.aleksic@ppf.edu.rs (M. Aleksić); jan.lang@stuba.sk (J. Lang)

🆔 0009-0001-6328-2004 (S. Milojković); 0000-0003-3538-6284 (P. Dakić); 0000-0002-0410-7724 (T. Heričko); 0000-0002-8675-7500 (M. Aleksić); 0000-0002-3271-7271 (J. Lang)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

of financial indicators. In this study, transparency refers to a clear link between input data and final interpretation, and accessibility means that the system can be used without prior technical or accounting expertise. The system is inspired by the notion that a small, carefully selected set of financial indicators such as current assets, short-term liabilities, and equity can provide sufficient insight into financial stability without requiring full data granularity [7].

At the core of the platform is the Ledger module, one of several other modules that present the structure of the balance sheet and generate step-by-step financial interpretations. This functionality builds on the principle of progressive disclosure [8] and follows recent advances in structured data processing [9]. Prior studies have shown that computational methods can reliably work with structured accounting inputs [10]. In particular, large language models have been applied to financial text and structured disclosures [11], and recent work has confirmed the effectiveness of XBRL formats for AI-driven financial analysis [12].

The main contribution of this paper is the design, implementation, and evaluation of the AI Ledger module within the ValidoAI system (see Section 4), which enables automated, explainable, and accessible balance sheet interpretation for SMEs using real accounting data. The Ledger module demonstrates how a minimal set of structured financial indicators can be transformed into narrative insights through AI-driven processing, supporting non-expert users in identifying key financial risks and making informed decisions without manual intervention.

The creation and preprocessing of the evaluation dataset are detailed in Section 4.1, which explains the step-by-step extraction, anonymization, and structuring of authentic SME accounting data. This rigorous process ensures that the data used for system validation is both representative of real-world accounting practices and fully compliant with privacy standards, laying a robust foundation for the subsequent analysis and interpretation presented in this study. The research process and methodology were guided by the application of the research questions listed within the section Section 2.2 Research Questions.

This paper is organized to guide the reader through a logical progression of ideas and findings as follows: Section 1 Introduction introduces the essential concepts and outlines the key challenges faced by small and medium-sized enterprises (SMEs) in navigating today's complex financial and regulatory environment. It also defines the main contributions of this research within the broader context of SME digital transformation.

Section 2 Materials and Methods surveys the existing literature, highlighting competing perspectives and emerging themes related to financial constraints, interpretive opacity, and regulatory friction that continue to affect SMEs.

Section 3 Literature Review presents the methodological framework of the study, combining qualitative and quantitative techniques. This approach is grounded in a real-world case study based on an authentic financial statement, provided with informed consent by a professional accounting agency client.

Section 4 Results focuses on the technical core of the project, detailing the architecture and operational flow of the Ledger module and the ValidoAI system. It explains how unstructured accounting data is processed into structured, actionable outputs through a layered workflow.

Section 5 Discussion analyzes the results in light of the research questions, discussing both the strengths and limitations of the proposed approach.

Finally, Section 6 Conclusion summarizes the main findings and outlines directions for future development, including data verification and broader system adaptation.

2. Materials and Methods

This study applies a structured mixed-methods approach that combines qualitative reasoning, quantitative data processing, and empirical evaluation. The methodological framework is based on a single case study conducted in a real-world setting, using actual financial documents provided by the Valido accounting agency.

To support transparency and reproducibility, the research team intends to release a portion of the anonymized dataset and preprocessing scripts under an open-source license (CC BY 4.0). All identifying elements, including company names, tax numbers, and account references, were removed or irreversibly hashed using the SHA-256 algorithm in accordance with data protection standards.

2.1. Dataset and Preprocessing Pipeline

The dataset consists of CSV exports derived from accounting software at the end of the 2023 fiscal year. It includes balance sheet data from two representative small businesses that were used as starting points in our research. Each file contains between 80 and 120 entries with columns such as item name, monetary value in RSD and EUR, share percentage, and a group classification label.

The data was processed using a Python-based ETL (Extract, Transform, Load) pipeline built with Pandas and NumPy. The process includes the following steps:

1. **Extraction:** CSV files were loaded automatically, without manual input or correction.
2. **Transformation:** Column headers were standardized, missing hierarchical values were forward-filled, and Serbian-style numbers (e.g., “1.234.567,89”) were converted to international decimal format (e.g., “1234567.89”). Non-analytical entries such as totals, section labels, and metadata were excluded.
3. **Classification:** Each entry was mapped to one of four core categories, defined in the balance sheet: fixed assets, current assets, equity, or liabilities, using rule-based keyword detection applied to the item descriptions.
4. **Loading:** The final structured dataset includes the following fields: item name, amount in RSD, amount in EUR, share percentage, financial category, and group label.

This structured format serves as the analytical foundation for generating financial indicators and supports subsequent narrative interpretation of the company’s balance sheet.

2.2. Research Questions

The methodology was shaped by the following research questions:

1. Can a small set of carefully selected indicators support reliable balance sheet interpretation for SME users without financial expertise?
2. Can structured financial logic and automated narrative generation produce interpretations that are consistent with standard accounting conventions?
3. What are the practical benefits and limitations of applying the MSQ methodology to automated financial analysis, particularly in terms of scalability, transparency, and SME-level usability?

3. Literature Review

This section, as an introductory part, reviews the main challenges SMEs face in financial data interpretation, summarizes recent advances in explainable AI and structured reporting, and discusses regulatory and security requirements relevant to automated financial analysis systems. Previous research and institutional reports by organizations such as the OECD, the World Bank, and the National Bank of Serbia emphasize the limited availability of analytical tools and the generally low level of digital literacy within the SME sector. At the same time, a gap persists between the complexity of advanced AI systems and their practical usability in resource-constrained environments.

A study on explainable AI (XAI) covered in this section with structured financial reporting underlines the importance of transparent and flexible solutions. The European regulatory environment requires high criteria for data protection, monitoring, and interpretability, creating further barriers to AI adoption in financial applications. Our results highlight the necessity for SME-specific systems that strike a compromise between technical efficiency, regulatory compliance, and practical application.

3.1. SME limitations in financial data interpretation

Reflecting on the literature, it becomes evident that SMEs face persistent barriers in developing financial capabilities and the right software solutions [13]. World Bank [14] highlights the lack of analytical tools and internal expertise as a fundamental obstacle to independent assessment of liquidity and solvency. This observation is echoed by OECD [2], who point to broader structural limitations, such as low financial literacy and limited digital engagement, which often render available tools ineffective. The regional perspective provided by the Regional Cooperation Council [4] further reinforces this view, noting a digital divide in the Western Balkans that continues to hinder the adoption of financial management software.

In considering the Serbian context, the National Bank of Serbia [15] identifies high transformation costs and a shortage of skilled personnel as key barriers to digital modernization, suggesting that awareness alone is insufficient for meaningful change. Bruegel [16] adds that institutional fragmentation and regulatory uncertainty further complicate technology adoption in transitional economies.

Fiagborlo and Kudo [17] observe that most accounting software for SMEs remains limited to static outputs, offering little interpretive guidance. Their findings, along with those of the OECD [18] and the World Bank [14], emphasize that high costs, weak infrastructure, and a lack of accessibility continue to impede effective use.

Recent advances in explainable AI offer a potential remedy. Liao and Varshney [9] argue that interpretability is not merely a technical feature but a user-centered necessity, with structured narrative outputs shown to improve comprehension for non-experts. Kim et al. [10] demonstrate that large language models can generate balance sheet summaries that align with human reasoning, thus supporting better decision-making. Çelik et al. [12] extend this by showing how structured XBRL data can be processed through AI models to produce narrative financial insights, thereby enhancing both accessibility and regulatory compliance.

Taken together, these studies suggest that the next generation of intelligent systems must move beyond automation to provide transparent, interpretable outputs that empower users—especially those without specialized financial backgrounds—to make informed decisions.

3.2. Financial data interpretation and structured reporting techniques

Kim et al. [10] examine the application of large language models (LLMs) in the context of financial disclosures. Their study demonstrates that such models can improve the consistency of interpretation, especially in environments where manual review is resource-intensive or unreliable. They argue that LLMs offer scalable solutions that can be embedded in automated financial analysis systems.

Çelik et al. [12] demonstrate how machine learning techniques can be effectively applied to XBRL-based financial data for tasks such as fraud detection and pattern recognition. Their findings support the view that structured formats like XBRL significantly enhance the ability of automated systems to generate consistent, interpretable, and regulation-aligned financial outputs, particularly in resource-constrained or manual-review-limited environments.

Liao and Varshney [9], from the perspective of explainable AI, emphasize the importance of transparency in algorithmic decision-making. They highlight that systems relying on narrative interpretations must be grounded in high-quality data and supported by clear, rule-based logic to be effective for end users.

These methodological principles are reflected in the ValidoAI system, which applies a structured, rule-based approach to the classification and interpretation of balance sheet items. By mapping data to predefined explanatory templates, the system generates consistent outputs that remain understandable to non-expert users. A detailed overview of the architecture and implementation is provided in the results section.

3.3. Regulatory and security perspectives in financial systems

European institutions have increasingly emphasized the importance of transparency, accountability, and data protection in the use of intelligent systems within the financial sector. The European Commission introduced the Artificial Intelligence Act (Regulation (EU) 2024/1689) [19], which outlines clear requirements for oversight, interpretability, and the ethical handling of sensitive data, as detailed in its 2023 policy framework. Similarly, the European Securities and Markets Authority (ESMA), in its 2025 guidelines [20], calls for financial systems that support auditability, human supervision, and predictable decision-making processes.

Research in this domain highlights that small and medium-sized enterprises face notable challenges in aligning with such standards. Limited technical resources and infrastructure often prevent SMEs from implementing features such as audit trails, data minimization techniques, and transparent decision logic. Within this context, system behavior that is both explainable and consistent becomes essential not only for meeting compliance requirements but also for maintaining user trust.

3.4. Security practices in AI-based financial systems

Security requirements for artificial intelligence in financial systems have become increasingly rigorous, especially concerning data protection and responsible automation. The European Commission, through the AI Act, together with the provisions of the General Data Protection Regulation, establishes key principles such as data minimization, auditability, and mandatory human oversight in automated decision-making [19].

These frameworks define not only legal obligations but also design expectations for AI systems intended for financial applications. ESMA [20], in its 2025 guidelines, reinforces this perspective by calling for transparent and accountable architectures, particularly in domains handling sensitive user information.

Kim et al. [10], in their analysis of large language models, raise concerns regarding the risk of unintended data retention and leakage when models are trained on non-anonymized datasets. Their findings emphasize the importance of controlling how data is processed and stored, particularly in models capable of producing human-like outputs that may inadvertently reveal private information.

This line of concern is echoed in studies that promote privacy-preserving design patterns, particularly those based on transient input processing. Researchers argue that such architectures, which avoid retaining user-level information, offer a practical compromise between interpretability and compliance. For SMEs, however, implementing these safeguards is often difficult due to constrained technical capacity and infrastructure, which further justifies the development of lightweight solutions that embed security principles without increasing complexity.

4. Results

The results presented in this section illustrate how the AI Ledger module within the ValidoAI system processes real-world accounting data and produces interpretable outputs for small and medium-sized enterprises. Emphasis is placed on the transformation of raw financial records into structured representations that reflect the company's financial position across key indicators such as liquidity, equity, and debt structure.

The following content is organized according to the system's operational flow, beginning with data ingestion and preprocessing, followed by categorization and anonymization steps, and concluding with the generation of narrative summaries. All outputs are visualized and contextualized to enable interpretation by non-expert users, ensuring practical applicability within resource-constrained business environments.

The system currently uses the gpt-4-0125-preview model via the OpenAI API, which requires internet access and is hosted as a cloud-based service. However, the architecture of the AI Ledger module is modular and does not rely on any model-specific features. This means that the narrative

generation component can be replaced with a smaller, local large language model (LLM), provided that it supports structured prompt handling and coherent financial interpretation. While the output quality may vary depending on the model used, the overall process is adaptable to different deployment environments, including on-premise setups.

The main limitation of this study is the potential lack of generalizability, as the evaluation is based on data from small and micro enterprises in Serbia. Local regulatory and infrastructural conditions may influence the applicability of results in other regional or economic contexts. Future research should involve a broader and more diverse sample, including cross-border data, to validate the approach across different environments.

4.1. Dataset and Preprocessing

In accordance with the presented previous content, we create the following items in data processing:

- 1. Dataset origin and content**

AI Ledger was developed to automatically process real-world financial reports generated by small and medium-sized enterprises using standard accounting software. The system was tested on both PDF and Excel documents without requiring manual formatting or data cleaning. Each file contained between 80 and 120 rows with fields such as RSD and EUR values, percentage shares, and item classifications. This structure allowed the system to operate consistently across firms and reporting periods, without human intervention.

- 2. Preprocessing pipeline**

To ensure numerical accuracy, a locale-aware parser was implemented to automatically detect Serbian-style number formatting (e.g., 1.234.567,89) and convert values into standard international decimal notation. This prevents parsing errors and enables consistent processing across data sources.

In addition, the system forward-fills missing hierarchical labels based on prior entries and standardizes column headers to ensure structural uniformity across documents.

This preprocessing stage serves as the foundation for all downstream analysis. It allows AI Ledger to process heterogeneous data inputs in a uniform way, without requiring manual adjustment. As a result, the system functions reliably even in real-world SME environments with minimal setup or technical supervision.

- 3. Balance sheet categorization**

Each entry in the report was automatically assigned to one of four core balance sheet categories: fixed assets, current assets, equity, or liabilities. This classification was performed using keyword detection within the item descriptions. For example, terms such as “equipment” and “buildings” were categorized as fixed assets, while “cash” and “receivables” were recognized as components of current assets.

This step established the structural basis for all subsequent interpretation performed by the system. By organizing data into logical categories, the module enables users to understand complex financial information through a familiar and accessible structure, even without formal financial training.

- 4. Data protection and reproducibility**

All identifying information, including company names, tax numbers, and account references, was removed or pseudonymized using the SHA-256 hashing algorithm. This approach ensured that the dataset retained its structural integrity while complying with data protection standards.

By applying this method, the AI Ledger system guarantees user privacy without compromising analytic value. Furthermore, the anonymized dataset and preprocessing scripts are prepared for public release on an open-access repository to support transparency, reproducibility, and further research.

- 5. Narrative prompt generation**

After the categorization process, the financial values were aggregated and transformed into a

structured natural language prompt. This prompt was submitted to the GPT model in order to generate a financial explanation that mimics expert-level reasoning.

The design of the prompt aimed to produce concise, context-specific output rather than generic AI text. The implementation logic behind this transformation is illustrated in Listing 1, which shows how categorized data is converted into an interpretable narrative structure.

A structured balance sheet with important financial categories that we previously described in the preceding paragraphs is displayed in Table 1, which was processed by the AI Ledger module. This includes property, plant, and equipment, which are examples of fixed assets. Thus, in this instance, we have: amount of 2,201 RSD (18,809.57 EUR), or 28.36% of the balance sheet overall. Current assets come next, which total 3,400 RSD (29,067.80 EUR), or 43.82%, and are mostly cash and bank accounts. Retained earnings-driven equity comes in second at 1,500 RSD (12,830 EUR), or 19.35%. At 8.47%, short-term liabilities total 650 RSD (5,560.50 EUR). Clear financial analysis is made possible by this breakdown, giving stakeholders a better understanding of how to use the system and plan their next course of action.

Table 1
Structured balance sheet entries processed by the AI Ledger module across key categories.

Code	Item Description	RSD	EUR	Share (%)	Category
A.	Fixed Assets	2,201.00	18,809.57	28.36	Fixed Assets
III.	Intangible Investments	–	0.00	–	Fixed Assets
IV.	Property, Plant, Equipment	2,201.00	18,809.57	28.36	Fixed Assets
B.	Cash and Bank Accounts	3,400.00	29,067.80	43.82	Current Assets
C.	Retained Earnings	1,500.00	12,830.00	19.35	Equity
D.	Short-term Liabilities	650.00	5,560.50	8.47	Liabilities

Using a GPT model, the given code snippet displayed on Listing 1 creates a financial story by processing categorized balance sheet data. It compiles information from a DataFrame, classifying it according to "Kategorija" (category) and adding up the values of "RSD" (Serbian Dinar) to create a dictionary (suma).

To ensure proper formatting with commas for readability, the prompt forms the essential balance sheet items of fixed assets, current assets, equity, and liabilities using the summed values. The OpenAI API is used to send this prompt to the GPT-4 model, which is defined as a senior financial advisor by a system role. For stakeholders, the model produces a concise, complete-sentence explanation of the balance sheet. This reasoning highlights the company's financial health and important KPIs while converting organized financial data into a logical story.

Listing 1: Core logic for generating a GPT-based financial narrative from categorized balance sheet data

```
1 suma = df.groupby("Kategorija")["RSD"].sum().to_dict()
2
3 prompt = f"""
4 Balance sheet as of 31.12.2024:
5
6 Fixed assets: {int(suma.get("Stalna imovina", 0))}, RSD
7 Current assets: {int(suma.get("Obrtna imovina", 0))}, RSD
8 Equity: {int(suma.get("Kapital", 0))}, RSD
9 Liabilities: {int(suma.get("Obaveze", 0))}, RSD
10
11 Please explain the balance sheet clearly and in full sentences.
12 """
13
14 response = openai.ChatCompletion.create(
15     model="gpt-4-0125-preview",
16     messages=[
```

```

17     {"role": "system", "content": "You are a senior financial advisor."},
18     {"role": "user", "content": prompt}
19 ]
20 )
21
22 text = response["choices"][0]["message"]["content"]

```

This code represents a key component of the AI Ledger module automated generation of narrative explanations based on raw financial data. It begins by aggregating all RSD (Serbian dinar) values across four main balance sheet categories: fixed assets, current assets, equity, and liabilities. These values are then embedded into a structured textual prompt, simulating a balance sheet summary as of a given date.

The prompt is submitted to a `gpt-4-0125-preview` model that has been configured to act as a senior financial advisor. Based on the categorized figures, the model generates a coherent narrative that explains the company's financial position in clear, connected sentences. This allows complex numerical reports to be transformed into accessible interpretations for users without formal financial training.

This method aims to translate technical accounting outputs into useful business insights, not only automate them. This allows for well-informed decision-making in settings with limited resources.

4.2. Interpretation with AI Ledger Module

The AI Ledger module converts raw balance sheet figures into clear insights that help users understand the company's actual financial position. Rather than presenting isolated amounts, the system generates narrative explanations based on categorized data such as current assets, short-term liabilities, and equity. These values are contextualized to indicate financial stability, highlight potential risks, and explain the relationship between assets and obligations. The output is written in accessible language, free of technical terminology.

This functionality is particularly valuable for entrepreneurs who lack access to financial analysts but need to make informed decisions. The technology allows for prompt action by interpreting data in addition to presenting it. In one instance, the algorithm identified a possible liquidity issue by detecting an imbalance between current assets and short-term liabilities. It provided a particular proposal rather than a general warning: investigate the working capital structure and think about extending the maturity of short-term loans.

In another case, a financial structure heavily reliant on short-term liabilities was identified as inefficient. The system suggested reallocating internal costs and reducing dependence on short-term financing. The strength of this module lies in its ability to transform technical data into actionable business insight. In resource-constrained environments, such interpretability provides essential support for managing financial risks effectively.

In addition, this mechanism is part of a broader functionality referred to as the AI Notebook, which offers users a unified narrative and visual representation of asset and liability structure. Through this component, users gain not only a technical interpretation but also a clear overview of the balance between equity, liabilities, and liquidity status. Automated comments highlight financial imbalances and propose concrete steps for improving the company's financial health.

To further assess the reliability of the selected indicators (as stated in RQ1), a manual validation was conducted using a representative case of a Serbian micro-enterprise. The balance sheet generated by the system was compared with the professional opinion of a certified accountant who was familiar with the firm's real financial standing.

4.3. AI Ledger System Architecture Overview

The architecture of the AI Ledger module is built around six sequential layers (Figure 1), forming a structured path from raw financial data to user-friendly interpretation. Each layer plays a specific role, and the overall design ensures transparency, scalability, and ease of extension.

The first layer ingests data directly from real-world business documents—bank statements, invoices, and payroll reports—without requiring manual formatting. The second layer immediately organizes this input by classifying each entry into one of the main balance sheet categories while also checking for inconsistencies.

After moving from structure to analysis, the third layer computes important metrics, including debt levels, equity share, and liquidity. The fourth layer “speaks” to the user by producing a narrative explanation that is clear and devoid of technical jargon, based on these measurements. The fifth layer suggests practical next steps, such as enhancing short-term solvency or reevaluating the liability structure, if imbalances are found. The sixth layer, which is available for use right now, provides the whole interpretation through an interface that blends informative text with visual charts.

This multi-layered method makes analysis more comprehensible rather than just automating it. AI Ledger provides structure, insight, and guidance to users who frequently function without professional assistance, but it does not take the role of a human accountant. Figure 1 shows the functional links between the processing layers of the system schematically. In this explanation, the system’s processing levels are shown in schematic form, with their functional relationships and the dynamic data flows between them elaborately mapped. The distinct roles that each layer plays in improving system effectiveness and smooth interoperability are carefully outlined. This tiered design does more than just automate analytical procedures; it turns intricacy into clarity, making the system’s functions both effective and easily understood.

Assets & Liabilities

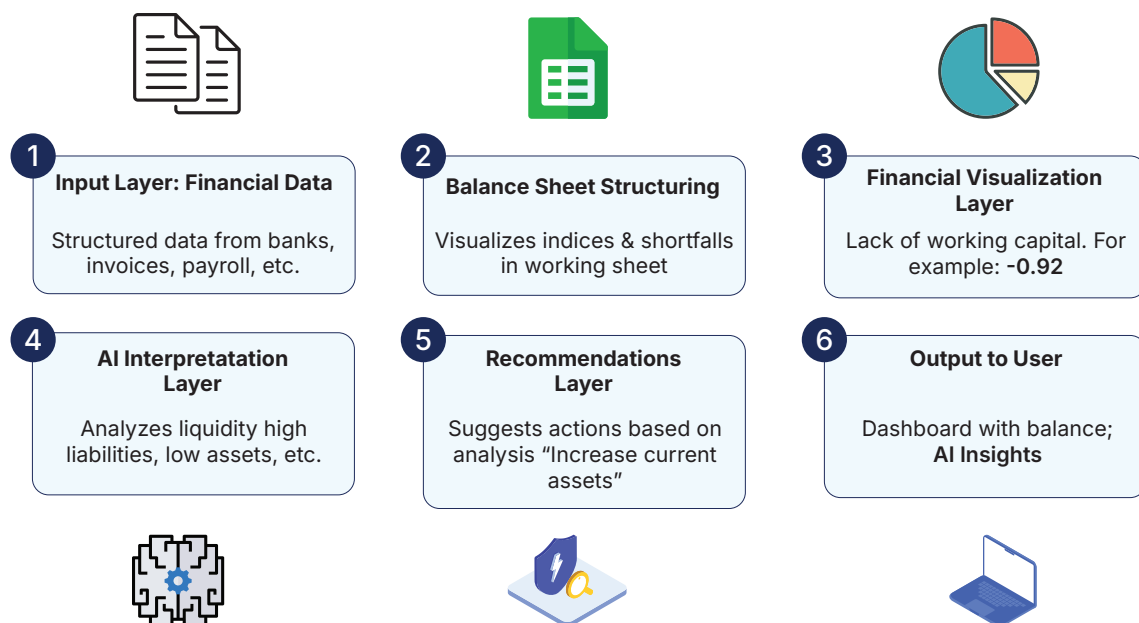


Figure 1: Modular structure of the AI Ledger system, showing six layers from structured data input to final user-facing interpretation.

In addition to the sequential data flow, the internal structure of the AI Ledger module is illustrated through a UML component diagram that shows the functional relationships among key modules. Figure 2 highlights three main segments: core processing, visualization, and AI-based interpretation. The core processing module includes components for preparing the balance sheet and converting Serbian-style number formats into standardized numerical values. Once processed, the data is forwarded to the visualization module, which generates visual outputs such as pie charts representing the asset-liability

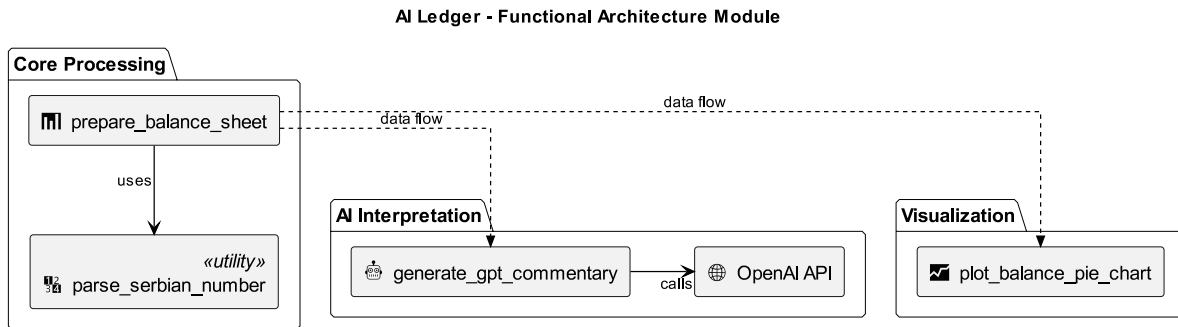


Figure 2: UML component diagram of the AI Ledger system.

structure.

The AI interpretation module takes the aggregated financial data and produces narrative commentary using a GPT model, triggered through an OpenAI API call. This allows the system to convert raw financial figures into explanatory text that is accessible to users without formal accounting expertise. Rather than serving as a purely technical overview, the diagram also illustrates the functional logic of the system from structured data transformation to clear, user-facing financial insights.

5. Discussion

The results confirmed that the AI Ledger module delivers accurate, explainable, and practically relevant interpretations of financial data using a limited set of structured inputs. The system successfully identified key relationships related to liquidity, capital structure, and liability distribution.

Unlike traditional dashboard systems that provide only numerical outputs, AI Ledger generates narrative explanations that clarify business implications. This enables users not only to detect problems but also to receive concrete recommendations without relying on expert interpretation.

The generated narrative outputs, grounded in structured financial logic, directly address the research questions and demonstrate the system's value in SME contexts with limited analytical resources.

Based on the above, we provide the following findings:

- RQ1.* With respect to the first research question—whether a limited set of financial indicators could support reliable balance sheet interpretation—the findings confirmed that four core indicators (fixed assets, current assets, equity, and liabilities) were sufficient for detecting key risk and stability signals.
- RQ2.* Regarding the second question, the GPT-4 model successfully generated narrative outputs that were aligned with accounting conventions and were easily interpretable by non-expert users.
- RQ3.* Finally, the third question, which addressed the practical advantages and limitations of the MSQ methodology, revealed that the approach was effective in constrained data environments but should be extended to include other financial dimensions such as cash flow and revenue segmentation in future iterations.

Despite promising results, small and medium-sized enterprises continue to face significant barriers to digital transformation, including financial costs, system complexity, and limited digital literacy. ValidoAI addresses these challenges through a simplified architecture based on minimal input and high interpretability, making it well-suited for environments with limited technical resources.

6. Conclusion

This study has shown that the AI Ledger module, created as a component of the ValidoAI system, offers a useful and workable way to automate balance sheet analysis, especially for small and micro-businesses. As we have discussed, the AI Ledger module has demonstrated its capacity to generate accurate insights by processing balance sheet data to produce a clear financial perspective. Property and equipment make up 28.36% of fixed assets, which total 2,201 RSD (€18,809.57). Cash and bank accounts make up the majority of current assets, which total 3,400 RSD (€29,067.80), or 43.82%. The equity from retained earnings is 19.35%, or 1,500 RSD (€12,830). The entire balance sheet showed 7,751 RSD (€66,237.87), while short-term liabilities totaled 650 RSD (€5,560.50), or 8.47%.

The results provide affirmative answers to all three research questions. They show that a small number of indicators can reliably assess liquidity and capital structure (RQ1), that the AI system produces explanations aligned with standard accounting logic (RQ2), and that the proposed MSQ-based approach is scalable, interpretable, and suitable for use in resource-constrained business environments (RQ3). The system relies on a minimal yet targeted set of financial indicators, including fixed assets, current assets, equity, and liabilities, to generate narrative insights from unstructured accounting data using GPT-4. Evaluation on real SME financial records confirmed that the model successfully interprets the company's financial position without requiring manual input.

It can be concluded that the main contribution of this work lies in the development of a lightweight and practical AI architecture. This empowers enterprises without dedicated financial expertise to independently evaluate their financial stability. This opens the door for broader adoption of AI tools in accounting and decision-making processes within the SME sector.

Future development will focus on expanding input data types, enabling multilingual support, and incorporating additional modules. This is important in the future for cash flow interpretation, VAT filing, and sectoral risk trends. The findings of this study demonstrate that well-designed, minimal-input AI solutions can play a critical role in future democratization of financial analysis for small and medium-sized businesses.

Acknowledgments

The work was supported by Erasmus+ ICM 2023 No. 2023-1-SK01-KA171-HED-000148295, the Slovak Research and Development Agency under Contract no. APVV-23-0408.

Declaration on Generative AI

During the preparation of this paper, the author(s) used Grammarly to do: Grammar and spelling check.

References

- [1] P. Dakić, L. Filipović, M. Starčević, Application of fundamental analysis in investment decision making: example of a domestic business entity, in: Selected Papers (part of ITEMA conference collection), ITEMA 2019, Association of Economists and Managers of the Balkans, Belgrade, Serbia, 2019, pp. 149–156. doi:10.31410/itema.s.p.2019.149.
- [2] OECD, Financial literacy and digitalisation for MSMEs in South East Europe, Technical Report, OECD, 2023. URL: <https://www.oecd.org/south-east-europe/financial-literacy-msmes-see.pdf>.

- [3] R. C. Council, Balkan barometer 2022: Business opinion survey, 2022. URL: <https://www.rcc.int/balkanbarometer/>, report.
- [4] R. C. Council, Western Balkans Digital Economy and Society Index (DESI) 2022 Report, Technical Report, RCC, 2022. URL: <https://www.rcc.int/docs/660/desi-2022-regional-report>.
- [5] T. E. Ali, F. I. Ali, P. Dakić, A. D. Zoltan, Trends, prospects, challenges, and security in the healthcare internet of things, *Computing* 107 (2024). doi:10.1007/s00607-024-01352-4.
- [6] P. Dakić, T. Heričko, v. Kljajić, V. Todorović, The economics of ai-powered call center development using chatgpt for the needs of an automotive retail business, in: *Proceedings of the International Scientific Conference - Sinteza 2025*, Sinteza 2025, Singidunum University, 2025, pp. 382–388. doi:10.15308/sinteza-2025-382-388.
- [7] J. Payne, Why bigger isn't always better where data is concerned, 2022. URL: <https://www.business-reporter.co.uk/digital-transformation/why-bigger-isnt-always-better-where-data-is-concerned>.
- [8] N. N. Group, Progressive disclosure helps users manage complexity, 2020. URL: <https://www.nngroup.com/articles/progressive-disclosure/>.
- [9] Q. Liao, K. Varshney, Human-centered explainable ai (xai): From algorithms to user experiences, 2022. URL: <https://arxiv.org/abs/2110.10790>.
- [10] A. Kim, M. Muhn, V. Nikolaev, Financial statement analysis with large language models, 2024. URL: <https://arxiv.org/abs/2402.11346>.
- [11] S. Wu, et al., Bloomberggpt: A large language model for finance, 2023. URL: <https://arxiv.org/abs/2303.17564>.
- [12] O. Çelik, V. İşler, M. A. Aydın, Financial statement fraud detection with machine learning techniques on xbrl data, *Applied Sciences* 13 (2023) 1597. URL: <https://www.mdpi.com/2076-3417/13/3/1597>. doi:10.3390/app13031597.
- [13] M. D. Haiderzai, P. Dakić, I. Stupavský, M. Aleksić, V. Todorović, Pattern shared vision refinement for enhancing collaboration and decision-making in government software projects, *Electronics* 14 (2025) 334. doi:10.3390/electronics14020334.
- [14] W. Bank, Fintech and SME Finance: Expanding Responsible Access, Technical Report, World Bank Group, 2022. URL: <https://documents1.worldbank.org/curated/en/099735004212257410/pdf/P17300606103440d8083ca064c26ca7f8dc.pdf>.
- [15] N. banka Srbije, Contemporary Data Sharing Models: Open Banking and Open Finance, Technical Report, Narodna banka Srbije, 2024. URL: https://nbs.rs/documents-eng/publikacije/wp_bulletin/wp_bulletin_09_24_2.pdf.
- [16] Bruegel, What is holding back artificial intelligence adoption in europe?, 2020. URL: <https://www.bruegel.org/policy-brief/what-holding-back-artificial-intelligence-adoption-europe>, policy Brief.
- [17] J. D. Fiagborlo, M. B. Kudo, Transforming accounting practices in small and medium-scale enterprises (smes): The roles and challenges of information and communication technology, *Journal of Money and Business* 4 (2024) 1–18. URL: <https://www.emerald.com/insight/content/doi/10.1108/jmb-09-2024-0054/full/html>. doi:10.1108/JMB-09-2024-0054.
- [18] OECD, Artificial Intelligence in Finance, Technical Report, OECD, 2024. URL: <https://www.oecd.org/en/topics/sub-issues/digital-finance/artificial-intelligence-in-finance.html>.
- [19] E. Commission, Artificial Intelligence: Economic Impact, Opportunities, Challenges, Technical Report, European Commission, 2023. URL: https://economy-finance.ec.europa.eu/document/download/10867a27-f49f-4aa9-89bb-feb752513f78_en?filename=dp210_en_artificial+intelligence_0.pdf.
- [20] E. Securities, M. A. (ESMA), Artificial Intelligence in EU Investment Funds, Technical Report, ESMA, 2025. URL: https://www.esma.europa.eu/sites/default/files/2025-02/ESMA50-43599798-9923_TRV_Article_Artificial_intelligence_in_EU_investment_funds.pdf.