

Article

Research on Optimization of M&A Financial Due Diligence Process Based on Data Analysis

Wei Li ^{1,*}¹ Advisory, Grant Thornton, New York, NY, 10017, United States

* Correspondence: Wei Li, Advisory, Grant Thornton, New York, NY, 10017, United States

Abstract: With the continuous development of data technology, enterprises have put forward higher requirements for the efficiency and accuracy of financial due diligence in merger and acquisition activities. The traditional due diligence process, due to many problems such as scattered data sources, lagging analysis, and insufficient security, can no longer meet the increasingly complex needs of mergers and acquisitions. This paper focuses on data analysis, conducts in-depth research from aspects such as the acquisition, processing and risk identification mechanism of financial data, proposes data-driven optimization strategies, builds a unified interface, real-time monitoring platform and data security system, which can enhance the intelligence, standardization and systematization level of financial due diligence in mergers and acquisitions, and provide more efficient and reliable support for enterprises' merger and acquisition decisions.

Keywords: data analysis; mergers and acquisitions; financial due diligence; process optimization; risk identification

1. Introduction

Against the background of economic globalization and the rapid development of digital information technology, enterprise mergers and acquisitions have become an important way of resource integration and scale expansion. Financial due diligence is a crucial component of the entire merger and acquisition process, significantly impacting the success or failure, as well as the risk control, of mergers and acquisitions. Currently, the due diligence process still faces significant challenges, including a high reliance on manual labor, low data processing efficiency, and lagging risk identification. With the continuous maturation of data analysis, intelligent modeling, and information systems, new possibilities have been provided for the optimization and intelligence of the due diligence process. This article takes data analysis as the core, explores and constructs an intelligent and efficient due diligence process model, promotes the transformation of financial investigation from traditional static investigation to dynamic monitoring, and enhances the scientific decision-making ability and risk management level of enterprises in merger and acquisition activities.

2. The Relationship between Data Analysis and Financial Due Diligence in Mergers and Acquisitions

2.1. The Importance of Financial Due Diligence in Mergers and Acquisitions

For any enterprise that intends to initiate a merger and acquisition, the financial due diligence of the merger and acquisition is a key means to understand the financial status, operational risks, and potential financial burdens of the target enterprise. Its core objective lies in revealing abnormal business practices and significant financial issues that may be hidden in the financial statements, such as low gross profit margin, excessive asset burden, and undisclosed contingent liabilities. Help enterprises determine reasonable merger and

Received: 21 August 2025
Revised: 04 September 2025
Accepted: 24 September 2025
Published: 27 September 2025



Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

acquisition prices, formulate corresponding risk prevention measures, and prevent valuation deviations or merger and acquisition failures caused by information opacity. Against the backdrop of an increasingly complex market environment and rising risk uncertainties, strengthening financial due diligence can enhance the success rate of mergers and acquisitions and ensure the safety of funds.

2.2. The Significance of Data Analysis for Financial Due Diligence

Financial due diligence can adopt data analysis techniques to efficiently collect and integrate multi-source financial data through automated means [1]. Then, after data cleaning and standardization processing, a structured data set suitable for modeling and analysis can be formed. By establishing a scientific analysis model, the system can quickly identify abnormal fluctuations and potential risks of key financial indicators, breaking through the traditional due diligence mode that relies on human experience for judgment. For merger and acquisition transactions, in the face of tight time and complex data, the adoption of data analysis can improve the investigation efficiency and is conducive to enhancing the forward-looking identification and judgment ability of potential risks of the target enterprise [2]. On this basis, by combining visual analysis and dynamic reporting mechanisms, the due diligence results can be presented more intuitively and transparently, enhancing the transparency and operability of financial information, providing a valuable analytical basis for investors, accelerating the merger and acquisition process, and promoting the smooth implementation of merger and acquisition transactions.

3. Construction of the Due Diligence Process Framework Based on Data Analysis

3.1. Process Architecture Design

The M&A financial due diligence process based on data analysis technology should comprehensively consider important links such as data collection, processing, and pre-processing, financial modeling, risk identification, and report output. The overall process framework is usually divided into five core working levels: data input, processing and calculation, risk analysis, decision support, and feedback optimization [3]. Each module is orderly connected based on data flow and logical relationships, forming a process architecture that is structurally reasonable, operationally efficient, and sustainably optimized. In practical applications, flowcharts can be used to clarify the dependency relationships and execution sequences among various functional modules, clearly display the indicator flow paths, the embedding positions of the system calculation models, and the response logic of the risk control mechanisms [4]. As shown in Figure 1, the entire system acquires information from multiple channels and operates. Then, it cleans, models, identifies, and outputs the information, and finally outputs it to the feedback mechanism, achieving the controllability, adjustability, and optimality of the due diligence process.

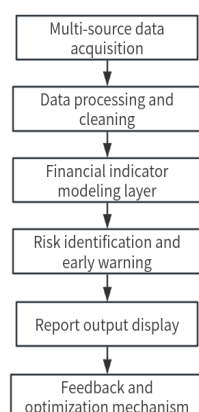


Figure 1. shows the structure diagram of the M&A financial due diligence process based on data analysis.

3.2. Intelligent Extraction and Processing of Financial Data

The core and prerequisite of the due diligence system based on data analysis is the intelligent extraction of financial data, which is the most crucial link in achieving automation and standardization throughout the entire process. The core task of this stage is to obtain structured data from various information platforms through technological means, and to clean, adjust, and improve the original data to ensure the uniformity and accuracy of the transmitted data.

The financial data of an enterprise mainly relies on ERP systems, financial statement management systems, and other third-party system platforms, among which there are also tax systems, industrial and commercial credit platforms, etc. By extracting key fields using a unified data interface and parsing semi-structured data and text data using the information extraction function $E(x)$, the data summary and integration from different application systems have been successfully achieved.

$$E(x) = \{f_1(x), f_2(x), \dots, f_n(x)\} \quad (1)$$

Among them, x is the original input data and is the extraction function of the field. Take "total assets" as an example. The system can automatically locate its position in different documents, extract the value and convert it into a standard unit (such as ten thousand yuan).

During the data cleaning process, the system will handle missing values, redundant items and format anomalies. The standardized function $S(x)$ is expressed as follows:

$$S(x) = Noemalize(x) + FillMissing(x) + ConvertUnit(x) \quad (2)$$

The above process unifies and integrates financial data such as balance sheets, income statements, and cash flow statements from different channels, providing a solid data foundation for subsequent model establishment and indicator comparison.

To ensure the accuracy of the data, after the extraction is completed, the system will automatically compare the historical data. If significant differences are found in the same field from different sources, it will be marked as an outlier ε , that is:

$$\varepsilon_i = \begin{cases} 1, & \text{if } |x_i^{(1)} - x_i^{(2)}| > \delta \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

Among them, $x_i^{(1)}$ and $x_i^{(2)}$ are the values of the same field from the two systems, respectively, and δ is the error tolerance threshold.

The intelligent extraction and processing of financial data not only enhances the efficiency and quality of data collection but also provide solid data support for model-based risk identification and analysis, laying a good foundation for the due diligence process.

3.3. Data-Driven and Indicator System Construction

In the M&A financial due diligence system based on data analysis, a scientific indicator system is the core foundation for achieving intelligent analysis. This system is based on assets, liabilities, profitability, capital stock and debt-paying ability, and constructs a multi-dimensional and multi-view indicator structure. The most fundamental dimensions include financial indicators such as the asset-liability ratio, current ratio, gross profit margin, and cash debt assumption ratio. According to the characteristics of the industry and the business model, composite indicators such as working capital efficiency, EBITDA growth rate, and cash conversion cycle can be introduced. In addition, integrating non-financial information such as supply chain stability and customer concentration can enhance the sensitivity and identification ability to external risks. To enhance intelligence and flexibility, the system can automatically select key features and adjust the index weights through machine learning, achieving dynamic updates of the index system. According to the characteristics of the industry and enterprises, the system can independently load the applicable indicator library and generate risk levels and early warning information through the risk scoring system, thereby promoting the transformation of

merger and acquisition decisions from traditional empirical judgment to data-based scientific decisions.

4. The Current Situation of the M&A Financial Due Diligence Process Based on Data Analysis

4.1. The Sources of Financial Data Are Scattered

During the financial due diligence process of mergers and acquisitions, the acquisition of financial data usually involves multiple internal information systems of the enterprise and external service interfaces, such as ERP systems, financial shared platforms, bank data interfaces, tax planning systems, and third-party data sources. Due to the non-uniformity of these interfaces in terms of universality and data structure standards, it often leads to situations of data redundancy, omission, or conflict. Furthermore, there are differences in data standards and dimensions among different departments, branches, and even accounting units, which further increases the difficulty of data integration. On the other hand, some data is still stored in unstructured forms, such as PDF reports or scanned vouchers, which cannot be directly analyzed. This reduces the efficiency of data processing and also affects the accuracy of model construction, becoming a bottleneck in the entire due diligence process.

4.2. The Monitoring of Key Financial Indicators Lags Behind

At present, in many M&A financial due diligence practices, there is a certain lag in the monitoring of key financial indicators, resulting in a slower speed of risk identification and lower accuracy of judgment. Most of the data comes from periodic reports, with lagging information updates, making it difficult to reflect the financial situation of enterprises in real time. Furthermore, a few enterprises may deliberately delay or tamper with the data submission time, increasing the judgment error. More importantly, the financial management system of enterprises does not yet have the functions of dynamic data collection and real-time early warning, and is unable to promptly identify important financial indicators such as abnormal cash flow and debt fluctuations, which reduces the risk prediction ability during the due diligence process.

4.3. The Security of Data Usage Is Low

During the financial due diligence process of mergers and acquisitions, the security of data is often overlooked, especially in the links of transmission, storage, and access, where potential security risks are prone to occur. Much core financial information, such as assets, profits, and tax records, is usually transmitted via email or shared networks. However, these transmission methods lack the necessary protective measures and are at risk of being stolen or tampered with. Furthermore, some companies have not established strict permission management systems, resulting in employees of different roles being able to easily access highly confidential information. Meanwhile, due to the lack of operation logs and behavior tracking, improper operations cannot be identified and corrected promptly. These problems not only threaten the confidentiality of the data but also may trigger legal disputes, thereby affecting the smooth progress of the acquisition transaction.

5. Optimization Strategy for M&A Financial Due Diligence Process Based on Data Analysis

5.1. Establish a Unified Data Interface to Integrate Multi-Source Information

To achieve efficient integration of multi-source data in financial due diligence, it is necessary to build a unified data interface platform. Through standardized data communication protocols and field mapping rules, the full-process data access to internal systems (such as ERP, financial shared service centers) and external platforms (such as industry databases, industrial and commercial tax systems) can be completed. The unified interface

can adopt the API integration mode, supplemented by ETL process scheduling, to achieve targeted extraction, transformation, and loading of data, and improve the automation and consistency level of data integration.

At the data structure level, the original system data D_s can be mapped to the unified data structure D_t required by the target analysis model through the mapping function M :

$$M: D_s, D_t = \{y_1, y_2, \dots, y_n\}, y_i = T_i(f_i(x_i)) \quad (4)$$

Among them, x_i represents the original field, f_i is the cleaning function, T_i is the format standardization function, and y_i is the integrated field. Each module completes the reconstruction of semantic consistency and structural consistency in the data platform through this transformation logic.

To ensure data quality and integration accuracy, the quality verification function $Q(x) = \alpha \cdot C_1(x) + \beta \cdot C_2(x) + \gamma \cdot C_3(x)$ can be introduced to quantitatively evaluate the accessed data from three dimensions: completeness, consistency, and accuracy.

$$Q(x) = \alpha \cdot C_1(x) + \beta \cdot C_2(x) + \gamma \cdot C_3(x) \quad (5)$$

Among them, $C_1(x)$, $C_2(x)$, and $C_3(x)$ respectively represent the verification scores of data integrity, logical consistency, and semantic accuracy, and $\alpha + \beta + \gamma = 1$ is an adjustable weight, which is used for flexible adaptation to different due diligence project scenarios.

In order to enhance the adaptability to multiple devices, interface components need to have meta-information-driven characteristics, be capable of real-time monitoring of field Settings and data modification rules in different systems, and support asynchronous or synchronous data transmission and version control. All input data should be connected to the data center. After undergoing homogenization processing and standardization correction, it should be directly transmitted to the risk identification module and the report generation module, thereby forming a complete numerical support system.

This method overcomes the problems of data segmentation and isolation, both inside and outside the enterprise. At the technical level, it ensures the integration, coordination, scheduling, and unified management of multi-source data, and serves as the basis for internal financial due diligence within the enterprise. This provides structured, high-quality, dynamic, and timely data support for intelligent analysis and modeling in the subsequent financial analysis and risk control stages.

5.2. Promote Real-Time Monitoring to Enhance the Timeliness of Indicators

In order to enhance the dynamic response capability of financial due diligence in mergers and acquisitions, it is necessary to introduce a real-time monitoring mechanism in the process to track and regulate key economic parameters in real time. Compared with the traditional way that relies on static reports, the real-time mechanism can identify abnormal changes more quickly, issue early warnings on time, and improve the timeliness and accuracy of risk judgment. The system collects key financial data such as sales profits, gross profits, asset-liability ratios, and accounts receivable through channels like enterprise ERP, bank interfaces, and tax systems. It automatically builds warning lines and sets update cycles, capable of achieving data updates every minute or every hour. This reporting model provides the review department with an advantage and consistency. Combining the set warning values and pop-up window functions, the system can visually display risk changes in the form of graphics or signals, helping reviewers intervene in a timely manner, thereby accelerating the review speed and improving the review quality. The following Table 1 shows some real-time monitoring financial indicators and their status that the system can support:

Table 1. Key financial indicators monitored in real time.

Indicator name	Current value (ten thousand yuan)	Daily variation rate(%)	Lower limit of the early warning threshold	Early warning status
Balance of accounts receivable	8,230.55	+4.3	-5.0	Normal
Operating income	12,945.80	-6.7	-5.0	Abnormal early warning
Inventory turnover rate	3.42	-0.2	-0.5	Normal
Operating cash flow	-1,120.00	-12.0	-8.0	Abnormal early warning

It can be seen from the table that the change in the enterprise's profit margin has exceeded the set warning value and entered the "abnormal warning" stage. The audit team needs to further analyze the abnormal situation of the enterprise's recent profit changes and cash inflows. Meanwhile, the accounts payable ratio and inventory turnover rate remain within the normal range, so no measures need to be taken. The early warning system effectively enhances the accuracy and response speed of due diligence.

5.3. Strengthen the Data Security Mechanism to Ensure Compliance with Usage

Merger and acquisition financial due diligence based on data analysis takes data as the core and utilizes a large amount of investigation data to assist in analysis and decision-making. In this mode, the security and legality of data are of vital importance. The investigation involves a lot of sensitive information, such as the company's asset structure, historical profit situation, tax records, and contract details etc. If these data are not handled properly, it may lead to problems such as data leakage, conflicts of interest, or legal disputes. Therefore, it is necessary to start from the entire process of data usage and establish a comprehensive security guarantee system from top-level design to the data end. Restrict the access rights of different investigators to financial data through role authorization. For example, in a certain manufacturing merger case, accountants can only view the summary part of the financial report of the target company, while the contract content is accessible only to lawyers. Through this authorization method, business secrets can be effectively protected, and the improper use of data can be avoided. In terms of data transmission, secure channels should be fully enabled, such as using TLS encryption when making API calls and downloading files, to ensure that sensitive data is not leaked during the transmission process. In a multinational enterprise merger and acquisition, due to the failure to activate the secure channel, the intermediate server was illegally accessed, resulting in the delay of the transaction. This incident has significant warning significance. During the data storage process, dynamic desensitization technology is adopted to process key data, such as company names, transaction amounts, and account information etc. In this way, the data accessed by relevant personnel in different scenarios is not complete. The system should automatically record operation logs to facilitate the monitoring and tracking of data usage, export, and modification.

6. Conclusion

This paper focuses on discussing the problems, such as data dispersion, lagging indicators, and weak security that occur in the financial due diligence of mergers and acquisitions, and establishes a process optimization framework based on data analysis. By introducing a unified data interface, real-time monitoring mechanism, and data security system, the efficiency of due diligence is enhanced. This proves that this work is of great

significance and wide applicability in the financial due diligence of mergers and acquisitions, and shows great potential in areas such as data-driven intelligent risk management and automated report generation.

References

1. P. Santoso, A. Purwanto, and M. Asbari, "Influence of Implementation Chain of Custody Forest Management System FSC-STD-40-004 V3-0 to Business Performance of Paper Industries in Banten Indonesia," *International Journal of Management and Humanities (IJMH)*, vol. 4, no. 4, pp. 32-36, 2019, doi: 10.35940/ijmh.D0482.124419.
2. C. A. Board, and C. Guard, "Federal Register," *Federal Register*, vol. 34, no. 100, 1969.
3. X. Yang, "Enterprise Financial Decision Support and Data Analysis," *Academic Journal of Business & Management*, vol. 6, no. 8, pp. 75-80, 2024, doi: 10.25236/AJBM.2024.060812.
4. Y. Li, and Y. Wu, "Research on optimizing teaching strategies for financial and accounting talents based on big data analysis," *Accounting, Auditing and Finance*, vol. 5, no. 1, pp. 37-44, 2024, doi: 10.23977/accaf.2024.050106.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of GBP and/or the editor(s). GBP and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.