ЕКОНОМІЧНЕ МОДЕЛЮВАННЯ
І ПРОГНОЗУВАННЯ

ECONOMIC MODELING
AND FORECASTING

https://doi.org/10.15407/economyukr.2025.06.039
UDC 658.012.011.56
JEL: C53, C81, M21

Viktor ZAMLYNSKYI, Dr. Sci. (Econ.), Prof.,
Professor of the Department of Industrial Economics
Odesa National University of Technology
112, Kanatna St., Odesa, 65039, Ukraine
Alla SHCHUROVSKA, PhD (Econ.), Associate Professor,
Associate Professor of the Department of Management and Marketing
State University of Intelligent Technologies and Telecommunications
1, Kuznechna St., Odesa, 65000, Ukraine
Fedir TRISHYN, PhD (Engineering), Associate Professor,
Vice-Rector for Scientific, Pedagogical, and Educational Work
Odesa National University of Technology
112, Kanatna St., Odesa, 65039, Ukraine

ANALYSIS OF DATA COLLECTION PROBLEMS IN INTEGRATED STRUCTURES AND THEIR IMPACT ON THE ACCURACY OF PREDICTIVE ANALYTICS IN BUILDING BUSINESS PROCESSES

The article analyzes how data collection issues in integrated structures affect predictive analytics accuracy. The need for standardized data collection, cleaning, and verification is emphasized. The role of machine learning and statistical methods, dependent on data quality, is considered. The conclusion highlights the importance of data control for improving analytical models.

***Keywords**: data collection; integrated structures; predictive analytics; data quality; machine learning; data integration into business processes.*

Predictive analytics plays a key role in intelligent decision-making in economics, business, and science. Its accuracy depends on the quality of the data used to build models. However, the data collection process can face issues that impact the outcome.

C i t a t i o n: Zamlynskyi, V., Shchurovska, A., Trishyn, F. (2025). Analysis of data collection problems in integrated structures and their impact on the accuracy of predictive analytics in building business processes. *Economy of Ukraine*. *68*. 06(763). 39-57. https://doi.org/10.15407/economyukr.2025.06.039

With the growth of digitalization and increasing data volumes, challenges arise related to the organization, storage, and processing of information, especially in integrated structures where data comes from multiple sources. Incorrect data collection, processing, or loss can lead to analytical errors that hinder operations.

The accuracy of predictions relies on the quality of the source data. Errors during the data collection phase can result in incorrect conclusions, losses, and strategic miscalculations. For example, errors in data regarding consumer preferences can lead to ineffective marketing strategies and inaccurate future forecasts.

As a result, ensuring the proper collection of data has become a critical issue for both scientific research and business practices. Modern data management methods, such as Data Governance, Data Quality Management, and the automation of data collection processes, are becoming increasingly important in the corporate environment.

Data collection methods: traditional data collection vs. automated, Web scraping, API and IoT-based data collection, integration of AI-based automation into data pipelines.

Data collection, analysis, and forecasting are crucial aspects of research in economics and business, especially in the context of improving the accuracy of forecasting models. In recent decades, scholars have focused on improving data quality, applying mathematical modeling, and using new information technologies to optimize analytical processes.

Modern data management techniques, such as Data Governance, Data Quality Management, and automated data collection processes, aim to minimize errors and enhance the reliability of analytical systems. However, despite significant progress, many issues remain unresolved, as evidenced by numerous studies from both international and domestic researchers.

For instance, *in Data Quality: The Field Guide,* T. Redman (2001) demonstrated that even minor errors in raw data could significantly distort analytical insights and forecasts. He emphasizes the necessity of stringent quality control at the earliest stages of data processing. These ideas were further developed by F. Provost and T. Fawcett (2013), who highlight how modern analytical approaches can be adapted for handling large datasets. They argue that integrating various data processing methods can mitigate individual weaknesses, and their research shows that systematically applying comprehensive data verification techniques significantly improves forecast accuracy, even when initial data contains minor errors.

Another critical issue is the influence of algorithms and data bias on predictive outcomes. C. O'Neil (2016) thoroughly examines how algorithmic bias can distort results and lead to flawed managerial decisions. Her work underscores the necessity of not only improving technical algorithms but also considering ethical aspects of data collection and processing. In this context, DJ Patil's research (Patil, 2014) demonstrates that analytics quality depends not only on methodologies but also on team efficiency. DJ Patil highlights the importance of interdisciplinary collaboration, which helps identify and correct data errors in a timely manner.

Furthermore, reliable infrastructures for data storage and processing play a key role in ensuring analytics quality. R. Kimball and M. Ross (2013) explore the construction of data warehouses and the organization of ETL (extract, transform, load) processes. His research shows that data structuring, format standardization, and process automation are fundamental to building accurate predictive models. Such an approach enables data integration from various sources and ensures its proper utilization in analytical systems.

The contributions of domestic researchers to the theoretical and practical aspects of data collection and analysis are substantial. V. Ventsel (2020) emphasizes the critical role of accurate data collection within integrated information systems, highlighting that analyzing large datasets requires both precision and adherence to established standards. He argues that errors at the initial stage of data collection can lead to significant inaccuracies in forecasting models, underscoring the necessity of specialized tools for data filtering and cleaning before processing. Ventsel's research further underscores the importance of organizing data collection within a unified system to minimize distortions and enhance forecast accuracy.

O. Shchedrina (2020) expands on this approach by addressing the challenges of integrating data from multiple sources through systems analysis. She highlights that insufficient data integration can result in inconsistencies, ultimately affecting the effectiveness of forecasting models. To mitigate information loss during data transfer across different levels and departments, O. Shchedrina advocates for the development of new methods to ensure data consistency and integrity.

S. Kravchenko et al. (2020) focuses on the technical aspects of data processing and classification algorithms in integrated environments. He emphasizes that selecting appropriate data processing algorithms — tailored to the structure and quality of the data — is essential for constructing accurate predictive models. His research suggests that leveraging machine learning techniques and automated data cleaning algorithms can significantly improve forecast accuracy while minimizing errors during the data collection phase.

Yu. Temchishina and I. Kravchenko (2014) complement these findings by examining the impact of data collection errors on business forecasting. They argue that inaccuracies in data can lead to flawed economic forecasts, which, in turn, negatively influence decision-making. Their work highlights the growing importance of data quality in the context of globalization and rapid environmental changes, where forecasting errors can result in considerable financial losses.

Collectively, these studies reinforce the crucial role of high-quality data collection and processing in ensuring the accuracy of predictive analytics. Challenges related to integration, data consistency, and preprocessing significantly influence forecasting outcomes and decision-making effectiveness. Therefore, the adoption of advanced data cleaning techniques, modern analytical algorithms, and robust integration frameworks is essential for improving the reliability of analytics within integrated systems.

Despite progress in the field, unresolved challenges remain, requiring further investigation. Data collection automation remains one of the most complex tasks in modern analytics, as companies must develop universal methods capable of accommodating diverse sources and dynamic business processes while minimizing errors at the initial stages of data processing. Additionally, ensuring high data quality necessitates the implementation of detailed validation protocols that minimize errors and eliminate duplicate records — prerequisites for building accurate predictive models.

Special attention should be given to data privacy and security, as increasing regulatory requirements demand a balanced approach between robust information protection and accessibility for analytical tasks. Moreover, ethical considerations, including mitigating algorithmic bias and minimizing the risk of discrimination, call for the establishment of appropriate standards. Furthermore, integrating structured and unstructured data remains a significant challenge, requiring new methods that leverage artificial intelligence and machine learning tools.

Based on a literature review, methodological approaches and practical recommendations have been formulated to improve data collection, cleaning, and management processes. Their implementation enhances analytics quality and forecast accuracy, ultimately enabling more informed managerial decision-making in a rapidly evolving economic landscape.

This study aims to provide a comprehensive analysis of the challenges associated with data collection and processing in companies, assessing their impact on analytical efficiency and predictive accuracy. The research explores aspects such as data inaccuracy, inconsistency, obsolescence, lack of standardization, and integration difficulties, analyzing their effects on analytical insights and managerial decision-making.

To achieve this objective, the study employs a systemic analysis approach to examine data collection, integration, and processing in complex information structures. A comparative analysis of predictive analytics approaches based on statistical methods, machine learning, and modeling is conducted. Additionally, mathematical modeling and factor analysis methods are used to assess the impact of data quality on forecast accuracy. The empirical section of the study is based on secondary data analysis, drawing from reports by leading analytical agencies and relevant scientific publications.

The data collection process in companies is complex and multifaceted, requiring adherence to numerous conditions to ensure high accuracy and reliability of information. The problem lies in the fact that companies often face:

- poor data quality (missing, duplicate, erroneous data);
- heterogeneity of information sources, making data integration more difficult;
- lack of standards and uniform data formats;
- access restrictions due to regulatory norms and privacy policies.

All these factors significantly hinder data analytics processing and reduce forecasting accuracy. Errors in data lead to incorrect management decisions, which can negatively affect the company's strategy and its competitiveness. Therefore, effective methods for data collection, cleaning, and management need to be developed to improve analytics quality and forecasting model accuracy.

The purpose of this article is to analyze the problems that arise in data collection and processing in companies and assess their impact on the effectiveness of analytical processes and the accuracy of forecasts.

The study conducts a comprehensive analysis of key issues related to data collection, such as inaccuracy, incorrectness, data obsolescence, lack of standardization, and integration challenges. It explores how these issues can influence the quality of analytical forecasts and decision-making processes' efficiency. In addition, the study includes an assessment of the impact of data quality on analytical and forecasting processes, focusing on changes in the volume of analytical information and the accuracy of forecasts across various fields. An important task of the research is also to develop recommendations for improving data collection and processing processes, which will allow for the identification of methodological approaches to data standardization and integration, as well as provide practical measures to enhance analytics quality and forecasting accuracy in companies.

This article analyzes the problems that arise during data collection in integrated structures and assesses their impact on the accuracy of forecasting analytics. The central focus is on how the organization of data within complex information systems of companies contributes to the emergence of errors, inaccuracies, and integration difficulties, which ultimately affect the quality of analytical conclusions and forecasting accuracy.

Forecasting analytics in modern business involves creating predictions of future events based on historical data analysis using statistical methods, mathematical modeling, and machine learning algorithms. The use of big data and complex IT systems within integrated structures not only allows for the collection and processing of information from various sources but also helps uncover hidden dependencies that affect the effectiveness of managerial decisions. However, it is in the integration of heterogeneous data where specific problems arise that need to be addressed to improve forecasting accuracy.

Historically, forecasting analytics began to take shape in the 1940s when A. Turing and his team (1950) applied mathematical methods to decode the German Enigma cipher machine. Over time, with the development of IT technologies, predictive analysis methods became widely applied in business environments, where data became the foundation for strategic planning. Today, companies aiming for sustainable development actively implement forecasting analytics systems that allow them to plan resource use and develop growth strategies.

An essential aspect of the study of data collection problems in integrated structures is the analysis of the impact of data quality on forecasting accuracy. Issues such as lack of standardization, inaccuracies, incorrect records, and difficulties in integrating heterogeneous information negatively affect the results of analytical processes. The forecasting analytics market in 2023 was valued at $14.71 billion, and it is projected to grow to $95 billion by 2032 [1] (Fig. 1). These figures highlight the increasing importance of quality data in forming reliable forecasts and making informed managerial decisions.

[1] Integration of Predictive Analytics with Traditional Business Intelligence (BI) Platforms to Aid Market Growth. *Fortune Business Insights*. 2023. URL: https://www.fortunebusinessinsights.com/predictive-analytics-market-105179 (accessed: 25.02.2025).

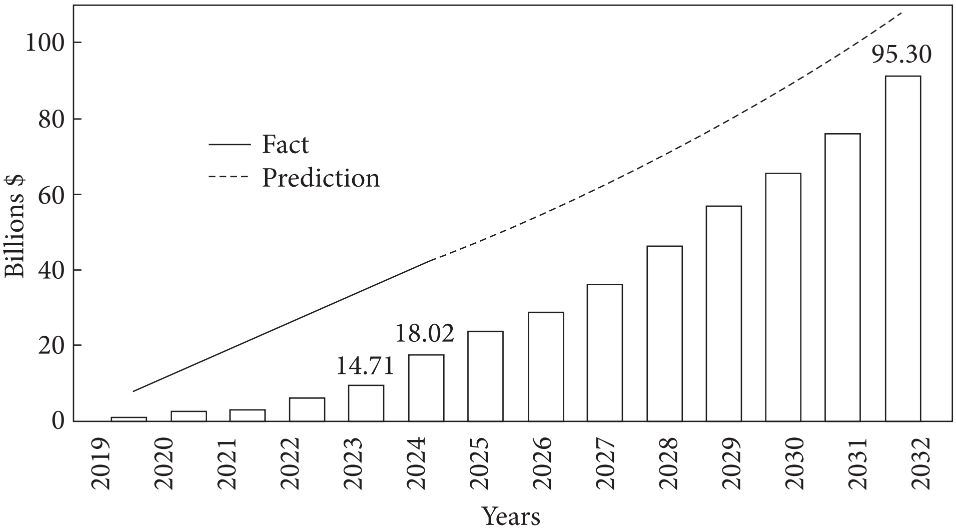

Fig. 1. Growth of the Forecasting Analytics Market from 2019 to 2032 Worldwide, billion $
Source: compiled by the authors based on: *Fortune Business Insights.* 2023. URL: https://www.fortunebusinessinsights.com/predictive-analytics-market-105179

Thus, it can be emphasized that the development of effective data collection and processing methods in integrated structures is essential to improve the accuracy of forecasting analytics. This, in turn, will enhance decision-making processes in the dynamically changing business environment and ensure the competitiveness of companies.

One of the key aspects of forecasting analytics is the selection of data processing methodology. Among the popular methods, the following can be highlighted (Fig. 2):

• **Regression:** this method allows assessing the relationship between variables and determining how a change in one variable (e.g., price) affects another (e.g., sales volume). Regression is used when processing large datasets and is a fundamental tool for quantitative analysis;

• **Decision Trees:** this method involves classifying data based on specified variables, which helps predict decisions made by individuals or the development of a particular process. The model is named for its resemblance to a tree structure, where branches represent possible choices, and leaves represent final outcomes;

• **Neural Networks:** based on machine learning principles, neural networks are used to model complex interactions between data when specific analytical formulas cannot be applied. These models are effective at handling tasks that require analyzing large numbers of factors and their interconnections.

The integration of data collected using the aforementioned methods into analytical systems significantly enhances the accuracy of forecasts and the quality of managerial decisions, especially in complex integrative structures. In the context of data integration within complex informational frameworks, predictive analytics represents a powerful tool for ensuring the objectivity of managerial decisions. The use of forecasting methods allows for quantitatively substantiated assessments rat-

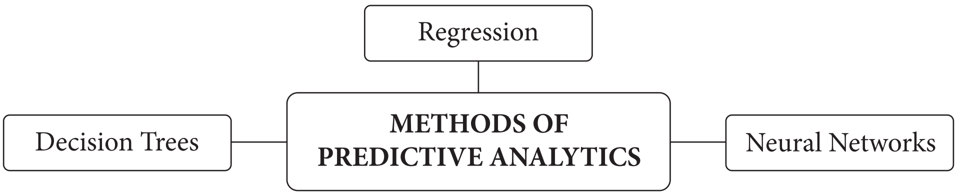

Fig. 2. Data Processing Methods
Source: compiled by the authors based on widely known data processing methodologies in predictive analytics.

her than relying solely on the subjective experience of managers, which helps reduce operational risks and optimize resource usage. For example, real-time pricing based on data allows for price adjustments according to changes in demand and market conditions, minimizing costs associated with inefficient decisions.

This area of analytics provides companies with the ability to quickly adapt to changes in the external environment. The more information available about future trends and events, the more flexible and competitive organizations become. Thanks to forecasting, companies can prepare strategic plans in advance, adjust product assortments, optimize supply chains, and thus gain an advantage over competitors.

Particular attention is paid to the impact of predictive analytics on financial business indicators. Forecasting demand changes, delivery times, and even potential equipment failures allows companies to adjust their operational processes in a timely manner. This not only improves operational efficiency but also increases profits through cost optimization and enhancing Customer Lifetime Value. Accurate forecasting models integrated into information systems support maintaining long-term customer relationships through timely and relevant marketing communications.

Predictive analytics is applied in a variety of business sectors, including retail, manufacturing, finance, logistics, healthcare, and telecommunications (Table 1). The Table 1 was compiled based on a synthesis of existing research and practical examples from various industries where predictive analytics is actively used. Information on the application of forecasting models in sectors such as retail, finance, marketing, manufacturing, healthcare, human resources, logistics, and cybersecurity was drawn from scientific publications and reports by leading analytical firms like McKinsey & Company, Deloitte, PwC, as well as industry research presented in open sources.

The data for the table was gathered through a literature review on predictive analytics, focusing on current trends and real-world examples of business implementation. It should be noted that the table reflects only the primary areas of predictive analytics application in these sectors and can be expanded or modified depending on the specific goals of the research and the available data.

Comprehensive data analysis allows for supply chain optimization, inventory management, improvement of pricing processes, and planning of marketing campaigns. The implementation of forecasting models contributes to increasing operational efficiency and sustainable business development, which is especially important in a rapidly changing business environment.

The basis of predictive analytics is the processing of large volumes of data collected from various sources in integrated information systems. The forecasting process can be divided into several key stages.

1. **Defining Goals:** at this stage, a specific business task is formulated, for which a predictive model needs to be built. An example might be the task of "determining the demand for product N in the next year".

2. **Choosing Forecasting Methods:** after the task is set, it is necessary to determine which analytical methods (regression, decision trees, neural networks, etc.) will most effectively address the goal, as well as which data sources and types of information will be used.

Table 1. **Application of Predictive analytics in Various Sectors**

Sector	How it is Applied
Sales	Demand and supply are assessed to accurately form product inventories. Competitive pricing is set, and options for cross-selling are developed based on customer preferences (using user data). In retail, demand changes are predicted based on seasonality and even climatic factors
Finance	In banking, risk forecasts are crucial for issuing loans. Machine learning algorithms analyze the likelihood that a customer will be able to repay the loan. Predictive analytics also helps combat fraud in the financial sector
Marketing	Creating offers that precisely meet customer needs. Data is used to create targeted advertising campaigns, i.e., personalization through various communication channels. This improves service quality by ensuring that unnecessary content and promotions are not presented to customers
Manufacturing	Predicting wear and tear and equipment failures to perform timely repairs. Also, forecasting production volumes and identifying factors that may affect the product being produced. Additionally, forecasting is crucial for workplace safety and preventing accidents
Healthcare	Using patient history and observations, potential disease progression scenarios are developed, and preventive measures are taken. Predictive analytics is also used in clinic operations: based on data, the required quantity of medications is ordered, and optimal scheduling is created
Human resources	With predictive analytics, employee alignment with company values is "forecasted", and the likelihood of employee turnover and its causes are calculated. Based on qualification data, lists of necessary skills for effective work are formed
Logistics	Forecasting is important for maintaining adequate inventory levels in warehouses and assessing situations to ensure timely deliveries
Cybersecurity	Potential ways of information theft are identified, allowing companies to minimize these risks or fully protect data storage. Cyberattacks are a major issue for companies, making forecasting in this area very important

Source: authors' design.

3. **Data Collection:** at this stage, data from various sources is integrated into a unified system (e.g., a data management platform or CDP). Depending on the field of analysis and the specific goals of the project, the necessary amount of information is selected, such as user behavior data, conversions, marketing activities, and others.

4. **Data Processing and Cleaning:** this is a crucial stage where unreliable, incomplete, or contradictory data is filtered out. Proper data cleaning is a key condition for building an accurate predictive model.

5. **Model Construction and Validation:** using the selected methods and cleaned data, a model is created, which then undergoes testing and adjustments. At this stage, the optimal set of parameters that allows the model to provide reliable forecasts is determined.

6. **Interpretation of Results and Implementation:** the final stage involves analyzing the obtained forecasts and integrating them into the decision-making process. The forecasting results serve as the basis for strategic planning, optimizing business processes, and adapting to changes in the external environment.

Thus, a comprehensive approach to data collection, processing, and analysis in integrated structures allows for significantly improving the accuracy of predictive analytics, but it also involves several risks related to the quality of the initial data and possible unexpected changes in the external environment.

Predictive analytics in integral structures provides significant advantages, but its implementation is associated with several risks that need to be considered.

First, even the most modern algorithms can make errors, and relying completely on predictions can lead to incorrect management decisions. *Second*, the effectiveness of analytical models directly depends on the completeness and accuracy of the input data. A lack of information or low-quality data can significantly reduce the accuracy of predictions, which will affect business management. *Third*, in rapidly changing external environments, there is always the risk of a "black swan" event — an unexpected occurrence that was not accounted for in the model and can render predictions invalid. For example, the COVID-19 pandemic radically altered business processes in many industries, despite all predictions. Thus, while predictive analytics simplifies decision-making, it cannot fully replace expert judgment and should be viewed as one tool among many to support management decisions.

An example of the negative consequences of poor data quality is the Gartner (2023) study[2], which found that 79% of strategic managers consider AI and analytics critical for success, yet only 16% of companies have fully implemented a data-driven approach. The primary reason cited for this is poor data quality, leading to inaccurate forecasts and diminished decision-making effectiveness.

Predictive analytics enables resource planning, risk minimization, and increased competitiveness. However, the success of predictive models directly depends on the quality of input data. Problems related to low-quality data, fragmented sources,

[2] Data-Driven Culture: Why Only 16% of Companies Fully Implement Analytics. Gartner. 2023. URL: https://www.gartner.com/en/newsroom/press-releases/2023-07-05-gartner-survey-finds-79-percent-of-corporate-strategists-see-ai-and-analytics-as-critical-to-their-success-over-the-next-two-years (accessed: 25.02.2025).

Table 2. **The Impact of Predictive Analytics on Company Performance, %**

Metric	Companies with Predictive Analytics	Companies without Predictive Analytics
Revenue Growth	+20	–2
Forecast Accuracy	85	50
Operational Efficiency	+30	–5

Source: compiled by the authors based on: McKinsey & Company. URL: https://www.mckinsey.com

and insufficient automation of data cleaning significantly reduce the accuracy of predictions and influence strategic decisions.

Main problems with data collection and processing in integral structures.

• **Low data quality:** duplicated or incorrect records in CRM systems, outdated information significantly reduce the accuracy of analytical models.

• **Fragmented data sources:** data from different systems often do not integrate into a single platform, making comprehensive analysis difficult and leading to the loss of important relationships.

• **Insufficient data volume:** to build reliable predictive models, a large amount of information is required. Limited data, for example, when analyzing short-term samples, may affect the objectivity of predictions.

• **Lack of automated data cleaning processes:** without an established procedure for filtering and cleaning data, unreliable information is used, which negatively affects analytics accuracy.

• **Human factor:** data entry errors (e.g., incorrect contact information or missing key fields) reduce the quality of input data and affect the reliability of predictions.

The analysis of the data collection process in integrated structures highlights the critical role of data quality on the accuracy of predictions. Issues such as lack of standardization, inaccuracies, incorrect records, and the complexity of integrating heterogeneous data adversely affect the results of analytical processes. For instance, according to the McKinsey report [3], companies that actively utilize predictive analytics and monitor data quality manage their operational processes 20% more effectively than their competitors.

Poor-quality data in integral structures lead to several negative consequences:

• reduced prediction accuracy;

• errors in initial information can lead to incorrect demand forecasting, causing issues with inventory management and production planning;

• wrong strategic decisions. Management relying on unreliable analytical reports risks making ineffective investment decisions, reducing the company's competitiveness;

[3] The Impact of Predictive Analytics on Business Performance. *McKinsey & Company.* 2023. URL: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year (accessed: 25.02.2025).

Table 3. **Key Data Quality Issues and their Impact on Forecast Accuracy, %**

Problem	Impact on Forecast Accuracy	Problem	Impact on Forecast Accuracy
Data Duplication	–15	Errors in Source Data	–25
Lack of Standardization	–20	Heterogeneous Data Sources	–18

Source: compiled by the authors based on: *Gartner*. 2023. URL: https://www.gartner.com/smarterwithgartner/how-to-improve-your-data-quality (accessed: 25.02.2025).

• financial losses. Investment in analytical systems without proper control over data quality leads to inefficient resource utilization;

• deterioration of customer experience. Inaccurate data can result in ineffective marketing strategies, reducing customer loyalty and conversion rates.

To improve the accuracy of predictive analytics, data processes need to be optimized. Key measures include:

• **automation of data collection and processing.** The use of ETL (Extract, Transform, Load) tools allows integrating and standardizing data from heterogeneous sources;

• **implementation of data quality management (DQM) systems.** These systems help detect and correct errors, remove duplicates, and maintain data integrity;

• **staff training.** Employees responsible for data collection and processing need to understand the importance of correct data entry and quality control;

• **regular data monitoring and updating.** Automated algorithms help maintain data relevance and accuracy.

The effectiveness of predictive analytics in integral structures directly depends on the quality of input data. Problems related to low quality, fragmentation, insufficient volume, lack of automated cleaning, and human error significantly reduce the accuracy of predictions and influence management decisions. Therefore, to successfully apply predictive models, companies must implement comprehensive measures to optimize data collection and processing, minimizing risks, improving analytics accuracy, and ensuring sustainable business growth (Zamlynskyi et al., 2024).

To ensure the accuracy of predictive models and minimize risks, it is necessary to improve the quality of input data, but also consider all aspects of transitioning to a data-driven approach in a company. Implementing such practices requires addressing a number of issues related to data management, which may be fragmented, outdated, or incomplete. One of the key steps is developing the infrastructure for systematic data processing and cleaning. Problems with data quality can significantly reduce analytical results and affect business strategy, emphasizing the importance of a comprehensive approach to these processes.

Modern companies are striving to implement a data-driven approach to improve product management and business processes. Big data analytics has become a key tool in strategic planning. However, transitioning to data-driven management is accompanied by several problems, the main ones being data quality and consistency. According to Fortune Business Insights, the global big data analytics market will grow from $348.21 billion in 2024 to $924.39 billion by 2032, with a CAGR of 13% [4]. According to Gartner's 2023 study, 79% of strategic managers consider AI and analytics critically important for future success [5]. However, only 16% of companies have fully integrated a data-driven approach into all business processes, due to data quality and source heterogeneity issues [6].

For a successful transition to data-driven management, it is necessary to understand the typology of business data: operational data, supply chain and partner data, customer data, financial data, personnel data, legal data, and performance metrics. In the absence of standardization and unified data collection protocols, these data may be stored in fragmented formats and subject to entry errors.

The transition of companies to data analytics is hindered by the following problems:

• lack of standardized data storage, heterogeneous formats, and system incompatibility;

• errors and gaps in data, such as duplication and outdated information;

• insufficient data quality control, lack of cleaning and validation processes;

• difficulty in integrating data from different sources, for example, difficulty merging information from CRM, ERP, and BI systems;

• limited data access due to legal restrictions and privacy issues.

According to VentureBeat 2019, only 13% of analytics projects reach production deployment, due to insufficient attention to data structure in the early stages [7].

To successfully transition to a data-driven approach, it is necessary to:

• automate data collection and processing by implementing ETL tools and standardizing formats;

• implement data quality management systems (DQM) that help detect and correct errors;

• train employees, increasing awareness of the importance of correct data entry;

• develop integration strategies using unified API standards and protocols to ensure data compatibility.

A flowchart for solving problems in the transition to a data-driven approach is shown in Fig. 3.

---

4 Big Data Analytics Market Size to Surpass USD 924.39 billion by 2032. *Fortune Business Insights.* URL: https://www.fortunebusinessinsights.com/big-data-analytics-market-106179 (accessed: 25.02.2025).

5 Gartner Dissects How AI and Analytics May Change Business Decision Making. *CX Today.* 2023. Jul 05. URL: https://www.cxtoday.com/customer-data-platform/gartner-dissects-how-ai-and-analytics-may-change-business-decision-making/#:~:text=A%20Gartner%20survey%20has%20found,and%20execution%20could%20be%20automated (accessed: 25.02.2025).

6 Gartner Survey Reveals Less Than Half of Data and Analytics Teams Effectively Provide Value to the Organization. *Gartner*. 2023. URL: https://www.gartner.com/en/newsroom/press-releases/03-21-2023-gartner-survey-reveals-less-than-half-of-data-and-analytics-teams-effectively-provide-value-to-the-organization (accessed: 25.02.2025).

7 Why do 87% of data science projects never make it into production? *VentureBeat.* 2019. Jul 19. URL: https://venturebeat.com/ai/why-do-87-of-data-science-projects-never-make-it-into-production/ (accessed: 25.02.2025).

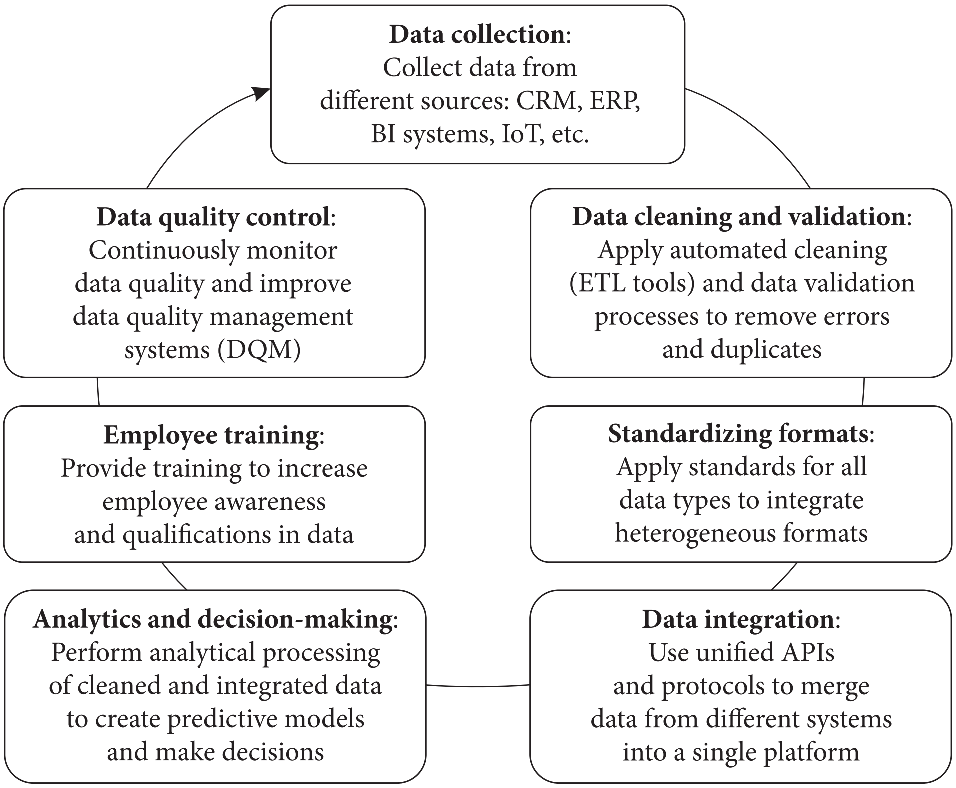

Fig. 3. Flowchart for Transitioning to a Data-Driven Approach
Sourse: compiled by the authors.

The analysis of data collection issues in integrated structures reveals that the success of a data-driven approach largely depends on the quality of the underlying information. The lack of a unified format, errors in manual data entry, and the fragmentation of data sources negatively impact the accuracy of predictive analytics. Therefore, companies must invest in standardization, integration, and data quality control today. This will not only improve the efficiency of management decisions but also significantly enhance the results of predictive analytics, which is a key factor for successful development in a rapidly changing market.

This study highlights the growing challenges of data collection in integrated structures, which can significantly impact the accuracy of predictive analytics in business process development. Empirical studies examining these issues focus on the application of marketing analytics in digital environments, including real-time price adjustments based on current demand and inventory levels, immediate fraud detection in transactions, market trend forecasting, and the promotion of innovation. These insights underscore the critical role of high-quality data management in optimizing decision-making processes and enhancing business adaptability, the need to "preparing a series of analytical reports in a management dashboard by analysing and mining all of the related project data" (Golestanizadeh et al., 2025). AI's potential to revolutionize marketing is discussed through analytical, technological, and strategic capabilities. Despite AI's transformative potential, its full impact across marketing roles remains untapped. V. Kumar et al. (2024) emphasize the importance of relevance, reliability, and relevance of data to ensure effective marketing activities. Insufficient data quality can lead to inaccurate forecasts and ineffective business decisions.

The next example is the development of accounting theory and methodology in the strategic management system. After all, the problems of collecting and processing statistical, accounting, and reporting data (determining research methods, the structure of data collection tools, and planning data collection) affect the accuracy of strategic forecasts and management decision-making. Incorrect organization of these stages can lead to inaccurate data, which will affect the quality of predictive analytics and the efficiency of business processes.

Our discussion centers on the emphasize the importance of high-quality data collection and processing in integrated structures to ensure the accuracy of predictive analytics and effective management of business processes.

Using data-driven analytics to assess the effectiveness of accounting for settlements with suppliers and contractors allows you to choose optimal solutions based on calculating their reputation for the entire period of activity and compliance with quality standards. This helps reduce resource costs and increase productivity. Business analytics allows you to diagnose and control costs, which will gradually lead to the modernization of accounting and reporting, changes in key financial indicators, analysis of personnel performance, providing continuous access to primary and consolidated accounting information.

The implementation of business analytics systems requires investment in digital infrastructure, analytical software and personnel training. A cost analysis should be conducted to assess the feasibility of using this technology and determine to what extent the use of business analytics can reduce operating costs and increase productivity. However, business analytics should not be considered a universal solution. The lack of data integrity, fragmentation of information systems make it difficult to create a comprehensive picture of the economic activity of the enterprise. Therefore, it is necessary to develop integrated databases and combine them with modern business analytics tools.

It is also important to implement a system of in-service training, which will increase labor efficiency and optimize the use of resources. Government agencies should provide funding for the creation of the necessary project management infrastructure, including quality control and information storage systems.

Among the main risks, it is worth highlighting the impossibility of predicting events caused by unpredictable factors. Recently, new approaches to predictive analytics have been developed to increase its adaptability. It is worth noting that these methods work better in developed countries with long-term strategies, but they should be adapted to modern challenges and global economic risks.

The development of integrated accounting and reporting systems contributes to the diagnosis of activities, assessment of the effectiveness of decisions and identification of risks. Transparency and speed of information processing increase the efficiency of management decision-making, which is a key factor in the development of modern business.

CONCLUSIONS

The results of this study confirm that the successful application of forecasting analytics in modern business structures is directly dependent on data quality and the effective integration of various information sources. Issues related to data heterogeneity, low quality, lack of standardization, and integration difficulties significantly reduce forecast accuracy and hinder the process of making informed management decisions. Based on these findings, the paper emphasizes the need for the standardization of data collection and processing, the implementation of automated data processing systems using ETL tools, the development and implementation of data quality management systems (DQM), as well as improving the qualifications of employees responsible for data handling.

The scientific novelty of the study lies in its identification of key factors that affect the accuracy of forecasting analytics in integrated information systems. Unlike previous research, this study not only focuses on data quality issues but also offers specific recommendations for improving these factors. In particular, new approaches to data standardization and the use of modern technologies for automating data processing and improving information quality are proposed, significantly enhancing the analytical outcomes. Additionally, the research underscores the need for a structured approach to data governance, which includes policies and procedures that regulate data management practices within organizations, ensuring consistency and reliability of analytical insights.

Furthermore, the study highlights the practical significance of continuous data quality monitoring, the implementation of data cleaning and validation processes to improve the accuracy of forecasting models, and reduce the risks of making incorrect management decisions. The scientific novelty also lies in the justification that poor data quality — such as duplication, format mismatches, and data entry errors — can significantly distort analytical results and reduce the effectiveness of forecasting models, which ultimately impacts strategic decision-making within a company. To address these issues, organizations must not only implement technical solutions but also cultivate a data-driven culture that prioritizes accuracy, consistency, and completeness of data.

The prospects for future research include the development of new methods for data integration, finding optimal solutions for standardizing information processing, and improving analytics quality within integrated structures. Specifically, it is important to explore the potential application of advanced technologies, such as artificial intelligence and machine learning, to improve forecast accuracy and create more flexible analytical models capable of adapting to rapidly changing market conditions. Further investigation into the role of big data, cloud computing, and real-time analytics in forecasting could also provide valuable insights into enhancing predictive capabilities. These measures will not only increase data-driven management efficiency but also strengthen organizations' competitive advantages by enabling more agile and precise decision-making.

Thus, for the successful implementation of forecasting analytics in business, it is necessary to consider not only technological aspects but also ensure a comprehensive approach to data quality management. This will significantly improve forecast accuracy, enhance decision-making processes, and ensure sustainable business development in the modern economy, where data is becoming a key resource for achieving competitive advantages. By integrating innovative technologies with robust data governance strategies, companies can maximize the potential of their data assets and establish a strong foundation for future growth and innovation.

REFERENCES

Redman, T. (2001). Data Quality: The Field Guide. Digital Press. 241 p. URL: https://books.google.com.ua/books/about/Data_Quality.html?id=iE2qsxDyEEYC&redir_esc=y

Provost, F., Fawcett, T. (2013). Data Science for Business. URL: https://www.researchgate.net/publication/256438799_Data_Science_for_Business

O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, Crown Publishers. 272 p. https://doi.org/10.5860/crl.78.3.403

Patil, DJ (2011). Building Data Science Teams. O'Reilly Media. 25 p.

Kimball, R., Ross, M. (2013). The Data Warehouse Toolkit. Third edition. Wiley. 601 p. URL: https://ia801609.us.archive.org/14/items/the-data-warehouse-toolkit-kimball/The%20Data%20Warehouse%20Toolkit%20-%20Kimball.pdf

Ventsel, V. (2020). Improving the effectiveness of management decisions in the process of economic and social policy-making based on economic and mathematical modeling. *Theory and Practice of Public Administration*, 1(68), 100-109. https://doi.org/10.34213/tp.20.01.13 [in Ukrainian].

Shchedrina, O. (2020). System analysis as a tool for making management decisions in business. *Modeling and Information System in Economics*, 99, 169-183. http://doi.org/10.33111/mise.99.15 [in Ukrainian].

Kravchenko, S., Grishkun, E., Vlasenko, O. (2020). Classification methods for machine learning using the Scikit-Learn library. *Scientific notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences.* Vol. 31 (70). No. 3. Part 1. https://doi.org/10.32838/TNU-2663-5941/2020.3-1/19 [in Ukrainian].

Temchyshyna, Yu., Kravchenko, I. (2014). Improvement of the accounting and analytical component in the information system for forecasting at a modern production enterprise. *Scientific Journal of the National Pedagogical Dragomanov University.* Iss. 26. P. 57-63. URL: https://dspace.nuft.edu.ua/handle/123456789/19632 [in Ukrainian].

Turing, A. (1950). Computing Machinery and Intelligence. *Mind.* Vol. LIX. Iss. 236. P. 433-460. https://doi.org/10.1093/mind/LIX.236.433

Zamlynskyi, V., Shchurovska, A., Zamlynska, O. (2023). Features and characteristics of business intelligence (BI)-systems as a tool for improving the efficiency of company activities. *Ukrainian Journal of Applied Economics and Technology.* Vol. 8. No. 1. P. 53-61. https://doi.org/10.36887/2415-8453-2023-1-8 [in Ukrainian].

Zamlynskyi, V., Shabatura, T., Zamlynska, O., Borysevych, E. (2024). Perspective Chapter: Exploring the Possibilities and Technologies of the Digital Agricultural Platform. Sustainable Development. Intech Open. https://doi.org/10.5772/intechopen.112358

Golestanizadeh, M., Sarvari, H., Parishani, A., Akindele, N., Edwards, D.J. (2025). Probing the Effect of Business Intelligence on the Performance of Construction Projects Through the Mediating Variable of Project Quality Management. *Buildings*, 15, 621. https://doi.org/10.3390/buildings15040621

Kumar, V., Ashraf, A., Nadeem, W. (2024). AI-powered marketing: What, where, and how? *International Journal of Information Management*, 77, 102783. https://doi.org/10.1016/j.ijinfomgt.2024.102783

Bousdekis, A., Kerasiotis, A., Kotsias, S., Theodoropoulou, G., Miaoulis, G., Ghazanfarpour, D. (2023). Modelling and Predictive Monitoring of Business Processes under Uncertainty with Reinforcement Learning. *Sensors*, 23 (15), 6931. https://doi.org/10.3390/s23156931

Received on February 07, 2025
Reviewed on February 21, 2025
Revised on March 03, 2025
Signed for printing on March 17, 2025

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

Redman, T. (2001). Data Quality: The Field Guide. Digital Press. 241 p. URL: https://books.google.com.ua/books/about/Data_Quality.html?id=iE2qsxDyEEYC&redir_esc=y

Provost, F., Fawcett, T. (2013). Data Science for Business. URL: https://www.researchgate.net/publication/256438799_Data_Science_for_Business

O'Neil, C. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York, Crown Publishers. 272 p. https://doi.org/10.5860/crl.78.3.403

Patil, DJ (2011). Building Data Science Teams. O'Reilly Media. 25 p.

Kimball, R., Ross, M. (2013). The Data Warehouse Toolkit. Third edition. Wiley. 601 p. URL: https://ia801609.us.archive.org/14/items/the-data-warehouse-toolkit-kimball/The%20Data%20Warehouse%20Toolkit%20-%20Kimball.pdf

Венцель, В. (2020). Підвищення ефективності управлінських рішень у процесі формування економічної та соціальної політики на основі економіко-математичного моделювання. *Теорія та практика державного управління.* Т. 1. № 68. С. 100—109. https://doi.org/10.34213/tp.20.01.13

Щедріна, О. (2020). Системний аналіз як інструмент прийняття управлінських рішень в бізнесі. *Моделювання та інформаційні системи в економіці.* № 99. С. 169—183. http://doi.org/10.33111/mise.99.15

Кравченко, С., Гришкун, Є., Власенко, О. (2020). Методи класифікації машинного навчання з використанням бібліотеки Scikit-Learn. *Вчені записки ТНУ імені В.І. Вернадського. Серія: технічні науки.* Т. 31 (70). № 3. Ч. 1. https://doi.org/10.32838/TNU-2663-5941/2020.3-1/19

Темчишина, Ю., Кравченко, І. (2014). Удосконалення обліково-аналітичної складової в інформаційній системі забезпечення прогнозування на сучасному виробничому підприємстві. *Науковий часопис НПУ імені М.П. Драгоманова.* Вип. 26. С. 57—63. URL: https://dspace.nuft.edu.ua/handle/123456789/19632

Turing, A. (1950). Computing Machinery and Intelligence. *Mind.* Vol. LIX. Iss. 236. P. 433—460. https://doi.org/10.1093/mind/LIX.236.433

Замлинський, В., Щуровська, А., Замлинська, О. (2023). Особливості та характеристики Business Intelligence (BI)-систем як інструменту підвищення ефективності діяльності компанії. *Український журнал прикладної економіки та техніки.* Т. 8. № 1. С. 53—61. https://doi.org/10.36887/2415-8453-2023-1-8

Zamlynskyi, V., Shabatura, T., Zamlynska, O., Borysevych, E. (2024). Perspective Chapter: Exploring the Possibilities and Technologies of the Digital Agricultural Platform. Sustainable Development. Intech Open. https://doi.org/10.5772/intechopen.112358

Golestanizadeh, M., Sarvari, H., Parishani, A., Akindele, N., Edwards, D.J. (2025). Probing the Effect of Business Intelligence on the Performance of Construction Projects Through the Mediating Variable of Project Quality Management. *Buildings*, 15, 621. https://doi.org/10.3390/buildings15040621

Kumar, V., Ashraf, A., Nadeem, W. (2024). AI-powered marketing: What, where, and how? *International Journal of Information Management*, 77, 102783. https://doi.org/10.1016/j.ijinfomgt.2024.102783

Bousdekis, A., Kerasiotis, A., Kotsias, S., Theodoropoulou, G., Miaoulis, G., Ghazanfarpour, D. (2023). Modelling and Predictive Monitoring of Business Processes under Uncertainty with Reinforcement Learning. *Sensors*, 23 (15), 6931. https://doi.org/10.3390/s23156931

Надійшла 07.02.2025
Прорецензована 21.02.2025
Доопрацьована 03.03.2025
Підписана до друку 17.03.2025

Віктор Замлинський, д-р екон. наук, проф.,
професор кафедри економіки промисловості
Одеський національний технологічний університет
вул. Канатна, 112, 65039, Одеса, Україна
Алла Щуровська, канд. екон. наук, доц.,
доцент кафедри менеджменту та маркетингу
Державний університет інтелектуальних технологій і зв’язку
вул. Кузнечна, 1, 65000, Одеса, Україна
Федір Трішин, канд. техн. наук, доц.,
проректор з науково-педагогічної та навчальної роботи
Одеський національний технологічний університет
вул. Канатна, 112, 65039, Одеса, Україна

АНАЛІЗ ПРОБЛЕМ ЗБИРАННЯ ДАНИХ У ІНТЕГРОВАНИХ СТРУКТУРАХ ТА ЇХ ВПЛИВ НА ТОЧНІСТЬ ПРОГНОЗНОЇ АНАЛІТИКИ ПРИ ПОБУДОВІ БІЗНЕС-ПРОЦЕСІВ

Ефективність бізнес-процесів у сучасному світі визначається якістю аналітики, проведеної на основі великих даних. Ці процеси являють собою послідовність взаємопов’язаних дій, спрямованих на досягнення стратегічних і оперативних цілей. Однією з ключових перешкод для високоточної аналітики є неоднорідність джерел даних, несумісність форматів, затримки в оновленні інформації, а також помилки в агрегації і обробці. Ці фактори провокують спотворення даних, що негативно впливає на надійність прогнозних моделей і може призвести до неефективного управління. Інструменти BigData дозволяють проводити детальний аналіз бізнес-операцій, виявляти «вузькі» місця, оптимізувати процеси і визначати перспективні напрями розвитку. Ефективність таких рішень безпосередньо залежить від якості вихідних даних. За наявності помилок, невідповідностей або суперечностей прогнозні моделі втрачають точність, що знижує достовірність аналітичних висновків. Для мінімізації ризиків, пов’язаних із перекрученням інформації, необхідно стандартизувати процедури збирання, перевірки й обробки даних. Використання сучасних алгоритмів машинного навчання і методів статистичного аналізу дозволяє автоматично виявляти помилки, очищати дані й підви-

щувати точність прогнозування. Завдяки оптимізації процесів збирання й аналізу даних у інтегрованих структурах компанії можуть ухвалювати обґрунтовані рішення на основі актуальної інформації. Це допомагає підвищувати конкурентоспроможність, створювати затребувані продукти, формулювати ефективні стратегії і швидко адаптуватися до змін на ринку.

Розглянуто вплив проблем збирання даних у інтегрованих структурах на точність прогнозної аналітики. Проаналізовано ключові фактори, що призводять до викривлення даних, включаючи неоднорідність, неповноту, застарілість і проблеми інтеграції. Обґрунтовано необхідність стандартизації процесів збирання, очищення і верифікації даних для підвищення достовірності прогнозу. Окреслено сучасні методи машинного навчання і статистичного аналізу, ефективність яких залежить від якості даних.

***Ключові слова**: збирання даних; інтегровані структури; прогнозна аналітика; якість даних; машинне навчання; інтеграція даних у бізнес-процеси.*