

ТИПОВЕ ДАННИ В УПРАВЛЕНИЕТО НА БИЗНЕС ПРОЦЕСИ (BPM) – КЛАСИФИКАЦИЯ, СРАВНЕНИЕ И АНАЛИТИЧЕН ПОТЕНЦИАЛ

Data Types in Business Process Management: classification, comparison, and analytical potential

Гено Стефанов¹,
e-mail: genostefanov@unwe.bg¹

Abstract

Business Process Management (BPM) generates and uses a variety of data types that play a key role in modeling, analyzing, optimizing, and monitoring business processes. This report provides a systematic classification and comparison of the main categories of process data used in BPM - process models (BPMN), process execution data (log data), and contextual business data. Their characteristics, sources, purposes, and applicability in various analytical scenarios are analyzed - from process mining to predictive analytics and business process optimization. A comparative summary of the advantages and limitations of each data type is presented, as well as their importance for generating knowledge and making management decisions. Finally, the main challenges related to their integration, quality, and joint use are outlined.

Key words: BPM, Process Mining, Predictive Analytics.

JEL: O30, O31

Introduction

In today's digital economy, Business Process Management (BPM) plays a crucial role in organizational sustainability, efficiency, and innovation[1]. A core feature of BPM is its data-centered nature — processes not only generate, but also extensively use various types of data, which form the foundation for modeling, monitoring, analysis, and continuous improvement. The quality of data and its preparation, including cleaning, transformation, and integration - is a critical prerequisite for the effectiveness of analytics in business process contexts[2]. In [3] the authors are emphasizing that data intelligence significantly influences how content is structured and presented in digital environments, which is increasingly relevant in the context of process - aware information systems. In [4] the authors are stating that digitalization has met certain expectations related to automation, efficiency, and accessibility, yet it has simultaneously exposed critical limitations - particularly concerning the integration of heterogeneous data and the realization of meaningful, data - driven business transformation, which directly impact how we use the process data.

Process data in BPM encompass different categories, each contributing uniquely to understanding and optimizing processes. This report provides a systematic overview of the three main categories of data in BPM:

- **Process Models** – created through notations such as BPMN, they formally describe the sequence, logic, and participants in a given process;
- **Process Execution Data (Logs)** – dynamic data recording the actual execution of processes within information systems such as ERP or BPM systems;
- **Contextual Business Data** – additional business data reflecting external and internal factors that influence process execution and analysis.

¹ Главен Асистент, доктор. Катедра Информационни технологии и комуникации, УНСС.

Each of these data types has a different structure, purpose, analytical potential, and set of challenges. Their integration and joint use are strategically important for knowledge extraction, precise analysis, and informed decision-making.

The goal of this report is to classify, compare, and analyze these three categories of process data, emphasizing their characteristics, applications, and role in modern analytical practices such as Process Mining, Predictive Analytics, and Business Process Optimization. Furthermore, the paper identifies challenges related to data quality, integration, and usability, which lay the foundation for future research and improvement in BPM.

Data categories in BPM

Process Models (BPMN)

Process models are formal and graphical descriptions of business activities, their sequence, participants, rules, and interrelations. They serve as a basic representation for understanding, analyzing, automating, and optimizing business processes. The most common notation is the Business Process Model and Notation (BPMN) — an international standard maintained by the Object Management Group (OMG)[1].

The main characteristics of this data type are listed in Table 1.

Table 2. Main characteristics of process model data

Characteristic	Description
Data Type	Structured (XML/JSON)
Format/Notation	BPMN, EPC (Event-driven Process Chains), UML Activity Diagrams
Components	Activities, events, branches, flows, roles, objects
Storage/Sources	BPMS (Business Process Management Systems), modeling tools (Camunda, Signavio, Bizagi)
Purpose	Modeling, documentation, automation, simulation
Level of Abstraction	High – describes what should happen, not necessarily how

Although process models primarily serve visualization and communication functions through graphical notations, their underlying structured digital representation — typically in XML or JSON — enables automated processing, analysis, and comparison with other process data sources.

Process Execution Data (Logs)

Process execution data, often called event logs or audit trails, are chronological records of actions, events, and states occurring during the actual execution of business processes within an information system.

They are automatically generated by various software systems such as BPMS, ERP (e.g., SAP, Oracle EBS), CRM, IoT platforms, and other applications that record interactions and transactions.

The main characteristics of this data type are listed in Table 2.

Table 2. Main characteristics of log data

Characteristic	Description
Data Type	Structured (tabular, JSON, CSV, XES – eXtensible Event Stream)
Granularity	High – records actions at the event level
Sources	BPMS, ERP, CRM, IoT devices, Web systems
Purpose	Execution analysis, Process Mining, monitoring, auditing
Issues	Different formats, lack of context

Logs contain structured information, most often using the following format (table 3):

Table 3. Example log data format

Поле	Пример
Case ID	INV 00231
Activity	Approve Invoice
Timestamp	10/29/2025 15:42
Resource	John Smith
Outcome	Approved
Context (opt.)	Amount > 1000, Country=BG

Logs are the main data source for Process Mining techniques, which include:

- Discovery – automatic extraction of process models from data;
- Conformance checking – comparing models with real executions;
- Enhancement – improving existing models with real-world data.

Contextual Business Data

Contextual business data enrich the understanding and analysis of business processes. Unlike process models and logs, which directly relate to process structure and execution, contextual data provide an “external perspective,” reflecting the environment in which processes occur.

They include both internal organizational factors (e.g., customer profiles, product categories, employee roles) and external influences (e.g., market conditions, seasonality, regulatory changes, macroeconomic indicators). The addition of contextual data significantly increases the analytical value of process analysis, especially in predictive and personalized scenarios.

The main characteristics of this data type are listed in Table 4.

Table 4. Main characteristics of contextual business data

Characteristic	Description
Data Type	Structured, semi-structured (JSON), unstructured (text, documents)
Sources	CRM systems, BI platforms, external databases, market research, open data
Purpose	Enriching analysis, forecasting, supporting decision-making
Issues	Heterogeneity, difficult integration, lack of standards, dynamic updates

The full potential of contextual data is realized when semantically linked to process events — for example, connecting a customer ID from a log to a CRM profile. This requires the use of metadata management, ontologies, and master data integration techniques.

Analytical potential and applications of data in BPM

Process Mining: From Execution Data to Real Insights

Process Mining is a data-driven approach that leverages event logs to:

- discover how processes are actually executed (discovery),
- compare real execution with predefined models (conformance checking),
- enhance existing models with insights from execution data (enhancement).

It is one of the most effective methods for identifying bottlenecks, deviations, inefficiencies, and opportunities for improvement[6].

Predictive Analytics: Anticipating Process Outcomes

By combining historical event logs with contextual business data, organizations can build predictive models that:

- estimate process completion time,
- forecast the likelihood of delays or failures,
- assess risk levels in ongoing process instances.

Techniques include machine learning algorithms such as Decision Trees, Random Forests, and Gradient Boosting, often enhanced by contextual features like customer segment, product type, or workload[10].

Prescriptive Analytics & Decision Support: Optimization through Insight

The highest level of analytics in BPM involves integrating all three data types to support[7]:

- scenario analysis and simulation (what if scenarios),
- process optimization (resource allocation, cost/time reduction),
- strategic decision-making via Decision Support Systems (DSS).

Figure 1. illustrates the intersection of three key data sources in Business Process Management: process models, execution logs, and contextual data. Each category contributes distinct value - process models provide structural definitions (e.g., BPMN), logs capture real execution behavior (e.g., XES), and contextual data enriches analysis with business relevance. At their intersection lie powerful capabilities such as process mining, predictive analytics, and simulation, ultimately enabling integrated BPM intelligence. This integration supports more informed decision - making, continuous improvement, and strategic alignment of business processes.

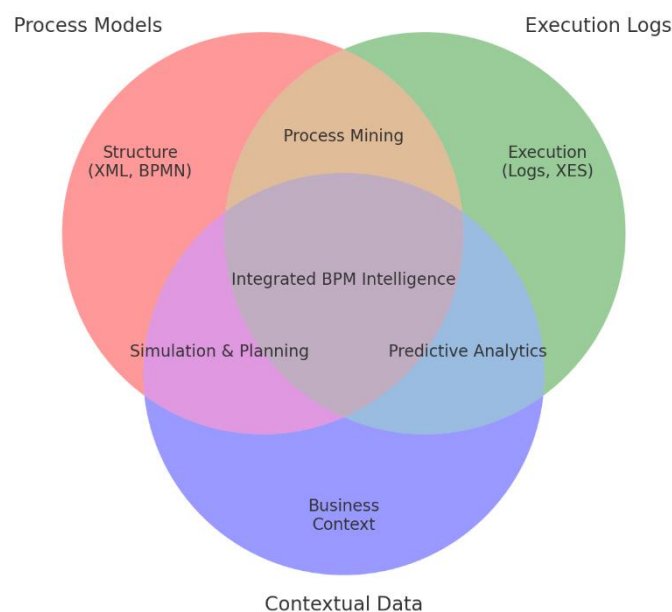


Figure 1: Intersection of BPM Data Types and Analytical use

Conclusion

The three types of BPM data — models, logs, and contextual data — represent complementary sources of knowledge. While each type has its own analytical potential, their integration and combined use yield the highest value for organizations. Major challenges remain: ensuring data quality, overcoming heterogeneity, and building compatible integration platforms. Future research should focus on developing integration frameworks, metadata standards, and AI-based methods for automatic linking between models and real data.

References

1. Belev, I., (2021). Comparing Business Process Management and Case Management, ICAICTSEE–2020, UNWE, Sofia, Bulgaria
2. Miteva, Genka. Murdzheva, A. (2023) Data Preparation Techniques and Platforms in the Context of Machine Learning <https://www.unwe.bg/doi/iited/2023/IITED.2023.21.pdf>
3. Велкова, Ивона. (2024) Намаляване на рекламните измами в дигиталния маркетинг с помощта на изкуствен интелект <https://www.unwe.bg/doi/iited/2023/IITED.2023.11.pdf>
4. Милев, П., Табов, Я. (2023) Влияние на интелигентността на данните върху структурирането на съдържание в веб публикации. Икономически и социални алтернативи, УНСС, бр. 4, 2023, с.73-85
5. Боянов, Л., Мурджева, А. (2024), Сбъднатите и несбъднатите очаквания от дигитализацията, <https://www.unwe.bg/doi/iited/2024/IITED.2024.15.pdf>
6. Comuzzi, M. (2024). Process Mining: From Descriptive to Predictive and Prescriptive Analysis. In Handbook on BPM.
7. Hasan, E. (2025). Big Data-Driven Business Process Optimization. TechRxiv.
8. Kokala, A. (2024). Leveraging Process Mining to Enhance BPM.
9. Sakr, S. et al. (2018). Business Process Analytics and Big Data Systems: A Roadmap. IEEE Access.
10. Márquez-Chamorro, A. et al. (2017). Predictive Monitoring of Business Processes: A Survey. IEEE Transactions.
11. Mihova, V., Murdjeva, A. (2014). Dynamic administration of database productivity. In Conferences of the department Informatics (No. 1, pp. 430-437).
12. Mihova, V., Marzovanova M. (2023) Modern Techniques and Tools for Generating SQL Queries, 13TH INTERNATIONAL CONFERENCE ON APPLICATION OF INFORMATION AND COMMUNICATION TECHNOLOGY AND STATISTICS IN ECONOMY AND EDUCATION (ICAICTSEE – 2023), UNWE, SOFIA, BULGARIA, <https://icaictsee.unwe.bg/past-conferences/ICAICTSEE-2023.pdf>