UDC 004.057.3+004.451.53+655.1

REFERENCE ARCHITECTURE FOR INDUSTRIAL INTERNET OF THINGS BIG DATA ANALYTICS IN SHORT-RUN PRINTING

T. V. Neroda[1]

[1]Institute of Printing Art and Media Technologies in Lviv Polytechnic National University, 19 Pid Holoskom St., Lviv, 79020, Ukraine;
ORCID: 0000-0002-5728-7060, e-mail: tetyana.v.neroda@lpnu.ua

Scientific research on engineering solutions for the integration of streaming and stored data flows remains insufficient, which complicates the implementation of end-to-end analytics, reduces production process transparency, and limits timely response to deviations in technological parameters. The relevance of this issue is emphasized, and the development of a parametric model for integrating operational printing production systems with target Big Data platforms for analytical processing of Industrial Internet of Things telemetry flows is justified. A multi-layer model of Big Data generation within a network of printing enterprises is structured, encompassing the communication layer of web-to-print order intake, the organizational layer of Customer Relationship Management and Enterprise Resource Planning information systems, and the technological layer of the production hub. The specific characteristics of file-based, event-based, and telemetry data flows at each layer are described, enabling the systematization of data sources according to their types and generation conditions.A mathematical formalization of production event flows is carried out, based on which the design of an event-driven reference architecture of a full-stack technological solution based on edge and cloud services is performed, eliminating fragmentation of the technology stack and ensuring an end-to-end cycle of analytical processing of heterogeneous production data.

Keywords: *short-run printing, Big Data, production metrics, event-driven layer, edge and cloud services, end-to-end analytics.*

Problem statement. Intensive digitalization of printing production, driven by the proliferation of network-integrated printing equipment, the development of software-controlled systems for data preparation and processing, and the implementation of web-oriented order intake and production planning services, leads to the rapid accumulation of large-scale and structurally heterogeneous datasets. In the domain of short-run printing, these data emerge in the form of streaming equipment telemetry, technological process logging, metadata of graphic file processing, and transactional records of management information systems. A key characteristic of such data is high ingestion rates combination, formats heterogeneity, and the requirement for targeted utilization to support mechanisms of streaming integration, asynchronous processing, and high-performance querying, while simultaneously ensuring production continuity and real-time decision support.

The accumulation of large volumes of heterogeneous data requires the application of integrated approaches to the organization of storage systems and processing workflows that enable the combination of traditional relational databases, analytical data warehouses, and caching systems with message streaming brokers and asynchronous processing tools. Such integration provides a unified technological pipeline that supports continuous ingestion, aggregation, and analysis of data from different types and sources. At the same time, it is necessary to consider the specific conditions of small and medium-sized printing enterprises, particularly limited computational resources and a minimal number of specialists capable of administering complex distributed systems, which determines the feasibility of using cost-efficient and interoperable software components.

Analysis of recent research and publications. The search for approaches to the development and improvement of Big Data processing methods across various domains and types of industrial organizations constitutes a significant research cluster in contemporary scientific studies, encompassing the analysis of existing models for processing streaming and stored data, their capabilities evaluation and limitations, as well as comparison of practical implementation outcomes in manufacturing and service environments.

In the production domain, multilayer stream processing models combine data aggregation, transformation, and real-time analytics, ensuring a consistent representation of heterogeneous data flows and improving the accuracy of detecting deviations in technological parameters [1]. Studies on the integration of sensor and operational data demonstrate the effectiveness of approaches based on distributed processing and adaptive resource management, which enable reduced latency and improved data consistency in complex production systems [2].

In analytical processing domain of stored data within business and service organizations, research efforts are focused on the development of multidimensional models for analyzing transactional, customer, and logistics data, as well as on integration of statistical and machine learning approaches for demand forecasting and risk assessment [3]. The proposed models enable the integration of structured and semi-structured data, supporting formation of adaptive analytical frameworks and improving decision-making efficiency. Studies further emphasize the applicability of hierarchical data organization approaches that provide scalability and flexibility of analytical processing under increasing data volumes, particularly in categorized processes of short-run printing production [4].

A separate research direction is associated with the development of Big Data processing models for small and medium-sized enterprises, where resource constraints determine the need for economically justified and technologically simple solutions [5, 6]. The cited studies propose hybrid architectural approaches that integrate asynchronous processing, multi-level caching, and adaptive data aggregation mechanisms, enabling the handling of heterogeneous information sources while maintaining a controlled level of system complexity [7].

Analysis of the reviewed publications indicates the presence of approaches aimed at reducing barriers to the adoption of Big Data technologies in resource-constrained enterprises, demonstrating significant interest in and demand for methods that enable a balance between data consistency, flexibility, and processing efficiency, thereby opening opportunities for further investigation of reference design approaches for full-stack solutions in small and medium-sized organizations. However, the issue of consistent integration of different data types [8] within a unified technological pipeline [9] remains insufficiently addressed. Existing approaches to the organization of information infrastructure in small business environments [10] are predominantly oriented toward the isolated use of individual software solutions performing accounting, file preparation, or equipment management functions [11]. Under such conditions, a unified model for the integration of streaming and stored data is absent, which complicates the implementation of end-to-end analytics, reduces the transparency of production processes, and limits the capability for timely response to deviations in technological parameters [12]. An additional factor is the use of diverse communication protocols and data representation formats, which complicates their unification and subsequent processing [13].

The situation is further constrained by the limited resources of small and medium-sized printing enterprises, which are characterized by restrictions in production capacity and computational infrastructure [9], absence of specialized support services for complex distributed systems, and the requirement to minimize costs associated with software deployment and operation [12]. Consequently, the adoption of full-scale enterprise Big Data processing platforms [13] is economically and organizationally impractical, whereas simplified solutions [14, 15] do not provide the required level of integration and analytical capabilities. Accordingly, there is a need for a holistic approach to the design of a reference architecture for Big Data processing that considers the specific characteristics of rapidly generated data sources throughout the order fulfillment lifecycle, ensures consistent integration of streaming and transactional data, and accounts for the resource constraints of small and medium-sized printing enterprises. The deployment of such an architectural solution should incorporate interoperable software components capable of supporting reliable data collection, storage, and processing across heterogeneous data types within a unified technological pipeline, with the capability for subsequent scaling.

Aim of article is to develop a parametric model for integrating short-run printing production systems with target Big Data platforms for the analytical processing of Industrial Internet of Things telemetry flows and to design an event-driven reference architecture of a full-stack technological solution based on edge and cloud services.

Presentation the main research material. Arrays of production telemetry nodes integrated with web-to-print platforms and telemetry flows form a multi-layer model of potential spontaneous Big Data generation within a network of printing enterprises. The structure of these data is characterized by high dynamics, heterogeneity, and a complex hierarchy of events and attributes, necessitating the formalization of data flows and parameterization for subsequent analytical processing and the design of reference architecture.

In the study of potential Big Data sources within networks of short-run printing enterprises, data flows are examined in relation to the order fulfillment processes during which they are generated. Since these processes belong to different segments of the printing-oriented infrastructure and involve the generation of heterogeneous data types, treating them as a homogeneous dataset complicates the identification of data generation and accumulation patterns. Therefore, for further analysis of Big Data sources, the production system should be represented as a structured model that associates data generation with corresponding process groups (Fig. 1). Such representation provides a basis for the organization and classification of data sources according to the types of data generated during print job execution and subsequent system operation.

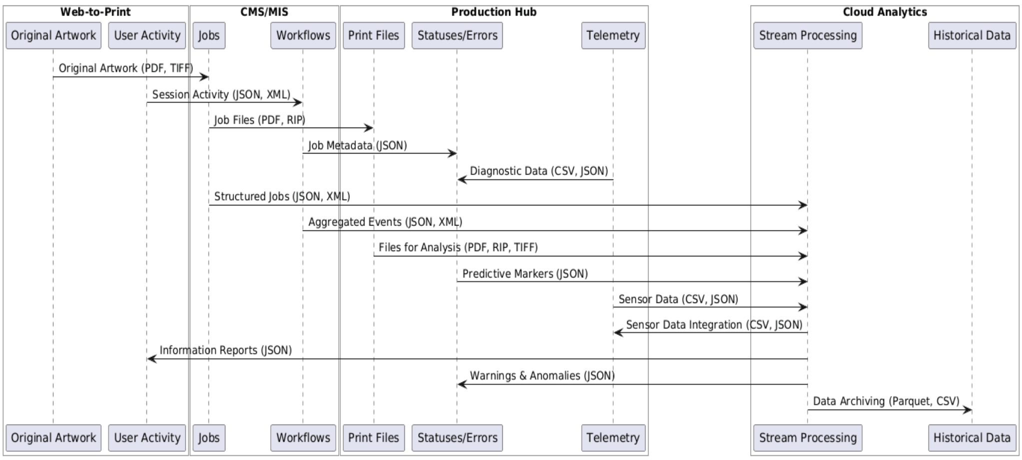

Fig. 1. Parametric model of workflow circulation in short-run printing production

At the communication layer of order intake (web-to-print), Big Data sources are represented by customer layout files, Portable Document Format documents, graphic templates, and media files uploaded through web platforms. Event-driven flows include incoming orders, order status changes, user sessions, Application Programming Interface requests, and access logs. Each of these elements forms an initial data flow transmitted to the organizational layer of order management systems. The volume and structure of such data depend on the intensity of user activity, the number of simultaneously submitted orders, and the diversity of file formats, determining potential peak loads and conditions for spontaneous Big Data generation.

At the organizational level of order management systems (CMS/MIS), data acquire a structured form and include print job files, production instructions, intermediate PDF files, and prepress processing files. Event flows comprise order creation and updates, workflow state transitions, and user and system activity logs. At this level, incoming flows from the web-to-print platform are aggregated and transformed, resulting in structured order and event flows transmitted to technological layer of the production hub. The structure of files and events at organizational level is characterized by increased variability and dynamic behavior, as workflow state changes and task modifications may occur asynchronously, leading to temporary accumulation of heterogeneous data arrays with varying structures.

The technological layer of the production hub constitutes the primary source of Big Data within the short-run printing production system. File-based data include RIP data, raster print images, and postpress files, including branded PDF files and layouts for folding and cutting operations. Event flows comprise print job execution statuses, equipment failures, and job queues within prepress and printing systems. In addition, this layer generates telemetry encompassing equipment performance indicators, diagnostic signals, process parameters, and machine sensor data. The concentration of event, file-based, and telemetry data at the technological layer determines the largest volume of information subject to analytical processing and creates conditions for the spontaneous emergence of Big Data flows during peak workloads or equipment anomalies.

Thus, event data are present at each layer of the production environment; however, their characteristics and structural complexity vary according to process specificity. Web-to-print layer is responsible for user-generated data and interaction events, organizational layer performs aggregation, transformation, and routing of data flows, whereas technological layer concentrates resulting files, production events, and equipment telemetry. This multilayer classification of Big Data sources provides basis for parameterizing data flows, developing integration models with Big Data platforms, and designing an event-driven reference architecture for analytical processing of production telemetry in short-run printing.

Event flow formalization in IIoT systems for short-run printing requires consideration of data source specificity, heterogeneity of structural representations, differences in generation frequency, and subsequent analytical utilization within the production environment. In contrast to traditional information systems, where data are predominantly transactional and accumulated in structured form, short-run printing environments are characterized by the concurrent operation of multiple independent event-generation domains, including network-integrated printing equipment, prepress graphic file processing systems, web-to-print services, production information systems, and operator actions associated with technological operations. Under these conditions, the establishment of a unified analytical environment requires prior formalization of event generation, transmission, and aggregation mechanisms describing the state of the production system.

In general, the short-run printing production environment can be represented as a set of data sources (1):

$$S = \{s_1, s_2, \dots, s_n\}, \tag{1}$$

where each element of the set s_i corresponds to a distinct class of production data sources. Such sources include digital printing presses, RIP-systems, order management software modules, web-based layout submission interfaces, warehouse material tracking systems, and sensor devices for monitoring physical equipment parameters. Representing data sources as a set enables a platform- and hardware-independent description of the information environment and supports subsequent scalability of the architecture.

Each source generates a stream of production events whose intensity and structure are determined by the characteristics of the technological process. The event stream for an individual source can be represented as (2):

$$E_i(t) = \{e_1, e_2, \dots, e_m\}, \tag{2}$$

where $E_i(t)$ denotes the set of events generated by source s_i within the time interval t, and each element e_j corresponds to an individual production event. Such events include print job initiation, completion of an order execution cycle, material feed errors, color correction parameter adjustments, completion of raster image processing, job queue state changes, or operator interventions in the printing process. Within a single production shift, the number of events may reach tens of thousands of records, while in a network of printing facilities they may form a near-continuous high-volume data stream.

A key characteristic of streaming data is arrival intensity, which determines the load on the information infrastructure and influences the selection of processing mechanisms. For an individual source, stream intensity is defined as the number of events generated by source s_i within a time interval Δt. This formulation captures heterogeneity across sources that produce data at different temporal granularities, ranging from equipment telemetry generated at sub-second intervals to events associated with the completion of discrete production operations or the creation of transactional records.

However, individual event flows (2) do not provide a complete representation of the production process, as they reflect only local states of individual subsystems. Consequently, there is a need to integrate heterogeneous event flows within a unified information pipeline. The aggregated production event flow can be represented as (3):

$$F(t) = \sum_{i=1}^{n} E_i(t), \tag{3}$$

where $F(t)$ denotes the integrated set of events received from all sources within the production system. This representation enables the fusion of equipment telemetry data, prepress file processing events, order lifecycle information, and administrative transactions, thereby providing an end-to-end representation of the print job execution lifecycle.

Nevertheless, the integrated data flow in its initial form (3) is not directly suitable for analytical processing due to significant heterogeneity, redundancy, and lack of contextual structuring. Consequently, the next step involves analytical event aggregation aimed at constructing structured representations of production processes. The generalized representation of this process can be expressed as (4):

$$A_k = f(F(t), p_k), \tag{4}$$

where A_k denotes the analytical representation of production data obtained by applying the transformation function f to the integrated event flow $F(t)$, subject to aggregation parameters p_k. Such parameters may include the observation time interval, equipment type, order identifier, category of technological operation, or production unit.

In practical implementation, aggregation involves transforming discrete events into structured analytical representations suitable for production state monitoring, statistical dependency modeling, and detection of deviations in technological parameters. Event sets associated with an individual order can be aggregated into temporal characteristics of file preparation stages, print execution duration, equipment stoppages, or job re-runs.

Similarly, telemetry streams can be grouped into time windows to evaluate equipment load levels, error frequency, and deviations from prescribed process parameters.

The formalization of event flows and their analytical aggregation enables a transition from fragmented accumulation of heterogeneous production records to a structured representation of short-run printing processes, establishing a foundation for subsequent design of an integrated analytical architecture for Big Data processing in the Industrial Internet of Things. Representation of the production hub as a set of data sources generating an integrated event stream highlights the asynchronous nature of data ingestion, heterogeneity of data formats, and the need to unify telemetry, transactional, and process-level data within a single information pipeline (Fig. 2). Under these conditions, the designed full-stack architecture must provide an event-driven interaction mechanism capable of supporting independent message ingestion from heterogeneous production sources without compromising processing integrity. Concurrently, the construction of analytical representations of production data requires separation of operational storage and historical data retention layers, necessitating distinct mechanisms for low-latency access and long-term analytical storage.

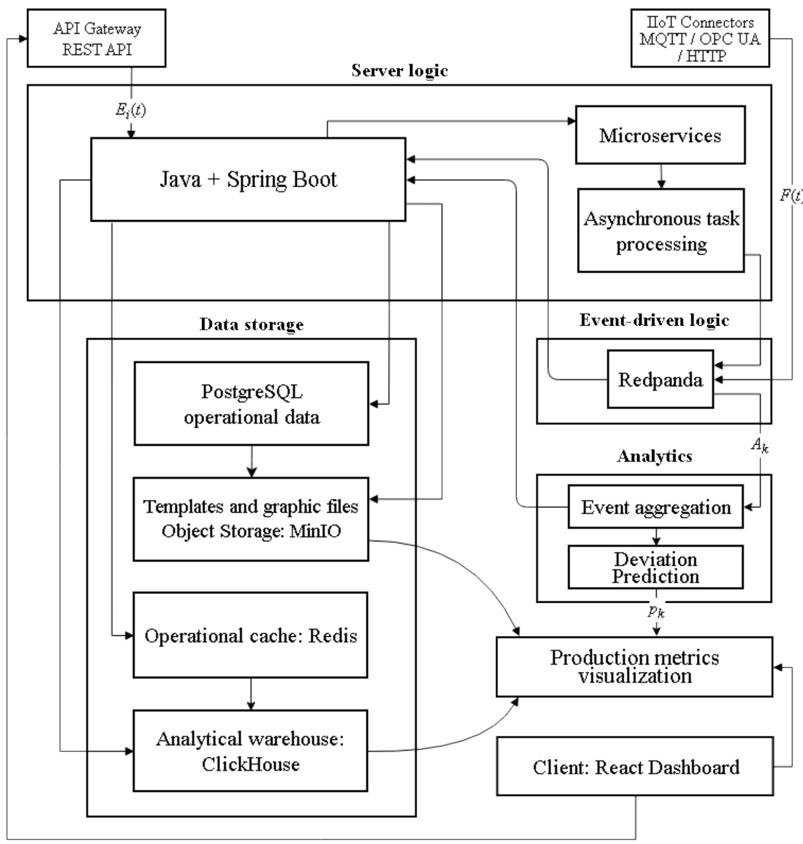

Fig. 2. Full-Stack Integration of the Production Hub with a Cloud-Based Analytical Layer

Heterogeneity of data sources necessitates support for multiple exchange formats and system scalability without altering the core logic of production data flow. Consequently, the proposed reference architecture for Industrial Internet of Things Big Data analytics in short-run printing constitutes a layered full-stack system built upon a unified event-driven data transmission and processing pipeline. At the data source layer, IIoT connectivity modules interfacing with production equipment via MQTT and OPC UA enable standardized acquisition of machine state telemetry, print parameters, process cycle events, and service signals (3). An additional data source is the Web-to-Print subsystem, which generates order lifecycle events and digital layout metadata (2). Event unification and delivery are implemented via Redpanda as a unified streaming broker acting as an ingestion layer, providing buffering, ordering, and routing of events while eliminating the need for multiple heterogeneous queues. All production and user-generated events are transformed into standardized messages and propagated through a unified streaming pipeline, thereby bridging operational and analytical processing layers.

The computational processing layer is implemented using a microservice architecture based on Java and Spring Boot. Services perform event-driven stream processing, including event filtering, normalization, and aggregation into intermediate analytical structures. Inter-service communication is realized via Redpanda, enabling an asynchronous data exchange model without dedicated message queues. The API Gateway provides a unified access point through REST interfaces for client applications and integration components. The data storage layer is separated into operational and analytical domains. Operational data are maintained in PostgreSQL as transactional backbone of production system. Redis is used for caching, providing low-latency access to order states and production queues. MinIO object storage, compatible with the S3 interface, is employed for storing templates, graphical assets, and large production workflow files. The analytical layer is implemented using ClickHouse, which accumulates aggregated events and supports near real-time analytical query execution. The analytical processing layer derives secondary indicators from aggregated event streams, including time series of production metrics, consolidated equipment load characteristics, and statistical estimates of process deviations (4). Processing results are delivered to visualization subsystem implemented as a React-based web client, providing real-time representation of production status and analytical indicators.

The proposed architecture constitutes a coherent system in which Redpanda serves as a unified event core, microservices provide distributed processing, and ClickHouse and MinIO implement the analytical and object storage layers, respectively. This design enables an end-to-end data processing pipeline from IIoT data sources to analytical representation, eliminating fragmentation across the technology stack.

Conclusions. The proposed target model for Industrial Internet of Things big data processing in small and medium-sized short-run printing environments integrates an event-driven paradigm for production data circulation with a multilayer structure for analytical processing. The architecture enables consistent transformation of technological equipment telemetry streams and order lifecycle events into a standardized analytical representation suitable for aggregation, anomaly prediction, and visualization of production metrics. Integration of edge and cloud computing layers provides functional separation of data ingestion, preprocessing, and analytical interpretation for high-velocity and structurally heterogeneous data streams. The resulting full-stack reference architecture supports incremental scaling of production capacity, is compatible with commonly used print-oriented infrastructures, and ensures interoperability required for system expansion, migration, and integration across heterogeneous production environments.

LIST OF USED SOURCES

1. Samadder M., Barman A. K., Roy A. K. Examining a generic streaming architecture for smart manufacturing's big data processing in anomaly detection: A review and a proposal // *International Journal of Experimental Research and Review*. 2023. Vol. 30. P. 219–227. doi: 10.52756/ijerr.2023.v30.019.
2. Bimonte S., Bellocchi G., Pinet F. et al. Data engineering for sustainable agriculture: Developments, challenges, and case studies of a novel IoRT architecture // *Journal of Big Data*. 2025. Vol. 12. Article 195. doi: 10.1186/s40537-025-01253-z.
3. Wang S., Wang H. Big data for small and medium-sized enterprises: A knowledge management model // *Journal of Knowledge Management*. 2020. Vol. 24, No. 4. P. 881–897. doi: 10.1108/JKM-02-2020-0081.
4. Neroda T. Big data categorization in operational digital printing processes // *Academic and University Science: Results and Prospect*. 2025. No. 18. P. 207–209.
5. Septian R. F., Abdillah F., Aliyudin T. F. A study review of common big data architecture for small-medium enterprise // *Proceedings of the 7th Mathematics, Science, and Computer Science Education International Seminar (MSCEIS 2019)*. EAI, 2020. DOI: 10.4108/eai.12-10-2019.2296535.
6. Cahyono D., Sijabat A., Panjaitan M. B., Julianingsih D., Eduaward A. L. Challenges and opportunities in implementing big data for small and medium enterprises (SMEs) // *Journal of Computer Science and Technology Application*. 2024. Vol. 2, No. 1. doi: 10.33050/corisinta.v2i1.74.
7. Ifthikhar S., Hamid B., Tahir S., Mehmood K., Sajid A., Banu S. Service-oriented architecture for big data analytics in smart cities // *Journal for Future Society and Education*. 2025. Vol. 2, No. 4. P. 20–42. doi: 10.3991/jfse.v2i4.59259.
8. Akidau T., Chernyak S., Lax R. *Streaming systems: The what, where, when, and how of large-scale data processing*. Sebastopol : O'Reilly Media, 2018. 349 p.
9. Neroda T. Analytical modeling of a dynamic discrete system for managing job-tasks in short-run print production // *Qualilogy Book*. 2025. Vol. 2, No. 48. P. 26–37. doi: 10.32403/2411-3611-2025-2-48-26-37.
10. Kusiak A. Smart manufacturing must embrace big data // *Nature*. 2017. Vol. 544. P. 23–25. doi: 10.1038/544023a.
11. Zaharia M. et al. Apache Spark: A unified engine for big data processing // *Communications of the ACM*. 2016. Vol. 59, No. 11. P. 56–65. doi: 10.1145/2934664.
12. Neroda T. Mobile application integration into the remote supervision subsystem of printing order preparation stages // *Computer Technologies of Printing*. 2025. Vol. 2, No. 54. P. 25–41. doi: 10.32403/2411-9210-2025-2-54-25-41.

13. Dubey R., Gunasekaran A., Childe S. J. et al. Big data analytics capability in supply chain agility // *Management Decision*. 2019. Vol. 57, No. 8. P. 2092–2112. doi: 10.1108/MD-01-2018-0119.
14. Wamba S. F., Gunasekaran A., Akter S. et al. Big data analytics and firm performance: Effects of dynamic capabilities // *Journal of Business Research*. 2017. Vol. 70. P. 356–365. doi: 10.1016/j.jbusres.2016.08.009.
15. Côrte-Real N., Oliveira T., Ruivo P. Assessing business value of big data analytics in European firms // *Journal of Business Research*. 2017. Vol. 70. P. 379–390. doi: 10.1016/j.jbusres.2016.08.011.

REFERENCES

1. Samadder, M., Barman, A. K., & Roy, A. K. (2023). Examining a generic streaming architecture for smart manufacturing's big data processing in anomaly detection: A review and a proposal. International Journal of Experimental Research and Review, 30, 219–227. doi: 10.52756/ijerr.2023.v30.019 (in Englich).
2. Bimonte, S., Bellocchi, G., Pinet, F., et al. (2025). Data engineering for sustainable agriculture: Developments, challenges, and case studies of a novel IoRT architecture. Journal of Big Data, 12, 195. doi: 10.1186/s40537-025-01253-z (in Englich).
3. Wang, S., & Wang, H. (2020). Big data for small and medium-sized enterprises: A knowledge management model. Journal of Knowledge Management, 24(4), 881–897. doi: 10.1108/JKM-02-2020-0081 (in Englich).
4. Neroda T. (2025) Big data categorization in operational digital printing processes. Academic and University Science: Results and Prospect, (18), 207–209 (in Englich).
5. Septian, R. F., Abdillah, F., & Aliyudin, T. F. (2020). A study review of common big data architecture for small-medium enterprise. In Proceedings of the 7th Mathematics, Science, and Computer Science Education International Seminar (MSCEIS 2019). EAI. doi: 10.4108/eai.12-10-2019.2296535 (in Englich).
6. Cahyono, D., Sijabat, A., Panjaitan, M. B., Julianingsih, D., & Eduaward, A. L. (2024). Challenges and opportunities in implementing big data for small and medium enterprises (SMEs). Journal of Computer Science and Technology Application, 2(1). doi: 10.33050/corisinta.v2i1.74 (in Englich).
7. Ifthikhar, S., Hamid, B., Tahir, S., Mehmood, K., Sajid, A., & Banu, S. (2025). Service-oriented architecture for big data analytics in smart cities. Journal for Future Society and Education, 2(4), 20–42. doi: 10.3991/jfse.v2i4.59259 (in Englich).
8. Akidau, T., Chernyak, S., & Lax, R. (2018). Streaming systems: The what, where, when, and how of large-scale data processing. O'Reilly Media, 349 p.
9. Neroda T. (2025). Analytical modeling of a dynamic discrete system for managing job-tasks in short-run print production. Book Qualilogy, 2 (48), 26–37. doi: 10.32403/2411-3611-2025-2-48-26-37 (in Englich).
10. Kusiak, A. (2020). Smart manufacturing must embrace big data. Nature 544, 23–25 (2017). doi: 10.1038/544023a (in Englich).
11. Zaharia, M., et al. (2016). Apache Spark: A unified engine for big data processing. Communications of the ACM, 59 (11), 56–65. doi: 10.1145/2934664 (in Englich).

12. Neroda T. (2025). Mobile application integration into the remote supervision subsystem of printing order preparation stages. Computer technologies of printing, 2(54), 25–41. doi: 10.32403/2411-9210-2025-2-54-25-41 (in Englich).
13. Dubey, R., Gunasekaran, A., Childe, S. J., et al. (2019). Big data analytics capability in supply chain agility. Management Decision, 57 (8), 2092-2112. doi: 10.1108/MD-01-2018-0119 (in Englich).
14. Wamba, S. F., Gunasekaran, A., Akter, S., et al. (2017). Big data analytics and firm performance: Effects of dynamic capabilities. Journal of Business Research, 70, 356–365. doi: 10.1016/j.jbusres.2016.08.009 (in Englich).
15. Côrte-Real, N., Oliveira, T., & Ruivo, P. (2017). Assessing business value of big data analytics in European firms. Journal of Business Research, 70, 379–390. doi: 10.1016/j.jbusres.2016.08.011 (in Englich).

doi: 10.32403/2411-3611-2026-1-49-206-217

РЕФЕРЕНСНА АРХІТЕКТУРА АНАЛІТИКИ ВЕЛИКИХ ДАНИХ ПРОМИСЛОВОГО ІНТЕРНЕТУ РЕЧЕЙ В ОПЕРАТИВНІЙ ПОЛІГРАФІЇ

Т. В. Нерода[1]

[1]Інститут поліграфії та медійних технологій Національного університету «Львівська політехніка», вул. Під Голоском, 19, Львів, 79020, Україна, ORCID: 0000-0002-5728-7060, e-mail: tetyana.v.neroda@lpnu.ua

Досліджено джерела стрімкого накопичення значних за обсягом і різнорідних за ускладненою структурою масивів даних в поліграфічно-орієнтованій виробничій інфраструктурі. Наголошено на важливості застосування комплексних підходів до накопичення та опрацювання даних, що дозволяють поєднувати традиційні реляційні бази даних, аналітичні сховища та системи кешування з потоковими брокерами повідомлень і інструментами асинхронної обробки. Виконано аналіз останніх досліджень та передових практик щодо узгодженої інтеграції різнотипових даних у межах єдиного технологічного контуру для організації інформаційної інфраструктури підприємств із обмеженими обчислювальними ресурсами, який продемонстрував відсутність узгодженої моделі інтеграції потокових і накопичених даних, що ускладнює реалізацію наскрізної аналітики, знижує прозорість виробничих процесів і обмежує можливості оперативного реагування на відхилення технологічних параметрів. Показано актуальність та обґрунтовано доцільність розробленні параметричної моделі інтеграції виробничих систем оперативної поліграфії з цільовими платформами Big Data для аналітичної обробки потоків телеметрії промислового інтернету речей Структуровано багаторівневу модель виникнення великих даних у мережі поліграфічних підприємств, яка охоплює комунікаційний рівень прийому замовлень, організаційний рівень систем керування та технологічний рівень виробничого хабу.

Охарактеризовано специфіку файлових, подієвих і телеметричних потоків для кожного рівня, що дозволило систематизувати джерела даних за типами й умовами їх генерації. Здійснено математичну формалізацію потоків виробничих подій, на основі чого виконано проєктування подієво-орієнтованої референсної архітектури повностекового технологічного рішення на основі межових та хмарних сервісів, що усуває фрагментацію технологічного стеку та забезпечує наскрізний цикл аналітичної обробки різнорідних виробничих даних.

Ключові слова: *оперативна поліграфія, великі дані, виробничі метрики, подієво-орієнтований контур, межові та хмарні сервіси, наскрізна аналітика.*

Стаття надійшла до редакції 04.04.2026.
Received 04.04.2026.
Прийнято до друку: 17.05.2026.
Accepted: 17.05.2026.
Опубліковано: 20.05.2026.
Published: 20.05.2026.