Journal of Information and Organizational Sciences
Volume 50, Number 1 (2026)
Journal homepage: jios.foi.hr

JIOS

DOI: 10.31341/jios.50.1.1
Open Access

UDC 004.031:005.591.6:336:338
Original Scientific Paper

Optimization of Batch Processing Through Advanced Management in Public Institution X

Mateja Gorenc[1*] and Janez Likozar[2]

[1]International School for Business and Social Studies, Celje, Slovenia
[2]B2 College of business sciences, Ljubljana, Slovenia
[*]Correspondence: mateja.gorenc@vspv.si

PAPER INFO

Paper history:
Received 09 September 2025
Accepted 30 December 2025

Citation:
Gorenc, M. & Likozar, J. (2026). Optimization of Batch Processing Through Advanced Management in Public Institution X. In Journal of Information and Organizational Sciences, vol. 50, no. 1, pp. 1-15

ABSTRACT

The article discusses the optimization of batch processing in a public institution through the implementation of a proprietary centralized batch management system. The research is based on the analysis of processing log entries from 2018 to 2023 and the implementation of a proprietary information interface developed in C# and connected to an Oracle database. The study highlights the importance of operator roles, structured work orders, and socio-technical alignment between technology and organizational processes. The analysis confirms that advanced planning significantly reduces processing time, thereby improving operational efficiency. However, the impact on overall process success is limited, as reliability appears to depend on additional organizational and infrastructural factors.

Keywords: batch processing, process optimization, log entries, automation, information systems

1. Introduction

In the modern information environment of public sector organizations, managing large volumes of data has become central to ensuring operational efficiency and the reliability of business processes. Particularly important are so-called batch processes, which form the backbone of the information infrastructure in many institutions where large datasets must be processed routinely and regularly. These are automated or manual computing procedures executed within predefined timeframes or triggered by specific events, with their success often depending on a complex network of technical, logical, and temporal dependencies.

Public institutions frequently face challenges such as fragmented execution of processes, limited centralized control, outdated workflows, and insufficient documentation, all of which increase operational risks and reduce efficiency. Although digitalization has improved many administrative functions, batch management in the public sector remains less advanced compared to private-sector environments. This gap is reinforced by differences in organizational constraints: public institutions typically operate within rigid legal frameworks, face resource limitations, and rely on legacy information systems that are not easily adaptable to modern optimization approaches. These factors influence not only the feasibility but also the impact of structured batch scheduling.

Despite the extensive research in financial and manufacturing industries, considerably less is known about how structured, centralized scheduling affects efficiency and reliability in public-sector systems. The novelty of this research lies in examining a real large-scale institutional environment over a six-year period and evaluating the effects of an internally developed centralized scheduling tool. The study combines socio-technical, contingency, and optimization perspectives to provide empirical evidence on how planning, dependency monitoring, and operator support influence batch process duration and success.

The goals of this research are threefold. First, we aim to quantify the impact of centralized and structured planning of batch processing on execution time and success rates in a real public-sector environment. Second, we seek to identify which aspects of planning and monitoring, such as dependency handling, automation of triggering, and operator oversight, contribute most to observed improvements. Third, we explore how insights obtained from log data can inform the design of future analytical tools and decision-support mechanisms. For IT managers and system administrators, the expected value of this work lies in concrete evidence on how a centralized scheduling tool can reduce delays and highlight remaining reliability issues. For institutional leadership and policymakers, the results offer a basis for investment decisions related to information infrastructure and human resources. For operators and technical staff, the study clarifies how structured work orders and clearer responsibilities can reduce operational uncertainty and error risk.

The research problem addressed is thus the following: *to what extent can centralized and structured management of batch processes improve duration and reliability in a public-sector institution, and what contextual factors determine the effectiveness of such interventions?*

Unlike studies that propose novel optimization algorithms or fully autonomous scheduling mechanisms, this paper does not aim to introduce a new scheduling paradigm. Its contribution lies instead in empirically evaluating the actual effects and limitations of centralized, rule-based batch planning in a real public-sector environment. By explicitly documenting both efficiency gains and their boundaries, the study seeks to provide a realistic and evidence-based understanding of centralized batch management, avoiding overly optimistic assumptions about automation as a universal solution.

The structure of the paper is as follows. Section 1 introduces the context, motivation, and research problem, and outlines the key challenges associated with batch processing in public-sector information systems. Section 2 presents the characteristics of batch processing and its relevance in the information systems of public institutions. Section 3 provides the theoretical framework of the study, which includes socio-technical approaches, the theory of constraints, and contingency theory. Section 4 formulates the research hypotheses. Section 5 describes the methodology, the analysed data, and the proprietary system for centralized batch management. Section 6 presents the results of the statistical analysis and hypothesis testing. Section 7 offers recommendations for improving existing processes. Section 8 concludes with the main findings and suggests possible directions for future research.

2. Batch Processing

Batch processing is a computational method in which a collection of tasks or application programs is executed sequentially without the need for direct user interaction. These tasks are typically carried out at specific time intervals, such as overnight or during periods of low system load, in order to optimize the use of computing resources and minimize disruptions to regular system operations. Batch processing thus represents the execution phase of an application program that operates based on predefined and clearly specified instructions (Thomas, 2024).

In financial institutions, batch processes often involve complex chains of tasks that handle large volumes of transactional data. For example, in payment processing, a batch workflow may include steps such as collecting and validating transaction data, posting to the general ledger, generating reports, and creating customer statements. Each step in this chain depends on the successful completion of the previous one, highlighting the importance of precise planning and control over the entire process (Khalfe, 2024).

Additionally, it is crucial to ensure that all data is properly stored and protected from loss or damage. This includes regularly creating data backups, especially before any changes or updates are made, to allow system restoration to a previous state in case of errors or failed processes. Doing so maintains data integrity and ensures the accuracy of outcomes, which is essential for user trust and regulatory compliance (Nicole, 2024).

Batch processing plays a critical role in the information systems of large organizations for several reasons:

- **Resource Optimization**: Batch processing enables the use of low-demand system periods to execute resource-intensive tasks, thereby avoiding system overload during peak hours and improving overall resource utilization efficiency (McHugh, 2024).
- **Automation of Repetitive Tasks**: By automating repetitive tasks, batch processing reduces the need for manual intervention, resulting in more consistent and reliable execution. This enhances productivity and minimizes the risk of human error (Talend, 2024).

- **Processing Large Volumes of Data**: Designed for the efficient handling of massive datasets, batch processing is particularly important for organizations dealing with extensive data, such as financial institutions, the healthcare sector, and government agencies. It allows for the processing and analysis of large data sets without straining system resources during peak periods (McHugh, 2024).
- **Reliability and Accuracy**: Batch processes ensure a high level of reliability and accuracy by executing tasks in a systematic and predictable manner, which is vital in business environments where data precision is critical (IBM, 2024).
- **Increased Operational Efficiency**: With centralized management and monitoring of batch processes, organizations can enhance operational efficiency. This includes tracking performance, optimizing processes, and responding swiftly to errors or failures. Centralized control enables better task organization and ensures alignment with business requirements (McHugh, 2024).
- **Integration with Various Systems**: Batch processing supports integration with different information systems and databases, simplifying complex business workflows and enhancing connectivity across multiple applications (McHugh, 2024).

Batch processing is a cornerstone of modern information systems, enabling automation, optimization, and reliable execution of numerous routine and data-intensive tasks. Through this, organizations achieve greater efficiency, reduced costs, and improved accuracy—factors essential for competitiveness and success in today's market (McHugh, 2024).

At the institution, we perform various types of batch processing:

- **Planned Processing**: These are pre-arranged processes executed according to a defined processing schedule.
- **Scheduled Processing**: These processes are not part of the main schedule but are automatically executed according to a predefined timetable, every day of the year, including weekends and holidays. Such processes ensure the continuous execution of critical operations, contributing to uninterrupted system functionality and compliance with operational requirements.
- **On-Demand Processing**: These are initiated at the request of the client, based on specific needs.
- **One-Time Processing**: These are processes that need to be executed only once, usually coordinated with a programmer or analyst.

The study builds on several theoretical perspectives that explain how structured planning and centralized management influence the performance of information systems. From the perspective of information systems theory, the socio-technical approach emphasizes the interdependence between technological solutions and organizational practices. In this view, batch processing efficiency cannot be achieved by technical improvements alone but requires alignment with organizational structures such as operator roles, work orders, and standardized procedures.

3. Theoretical Framework

The theoretical foundation of this study is based on socio-technical systems theory (STS), which emphasizes that the effectiveness of information systems depends on the alignment between technical solutions and organizational as well as human factors. As early as Trist and Bamforth (1951), it was demonstrated that technological improvements alone do not lead to greater efficiency if they are not simultaneously adapted to work practices and the social dynamics of the organization. Mumford (2006) highlighted the importance of participatory design, which fosters greater acceptance of information systems, while Alter (2015), through the development of Work System Theory, extended STS by offering a more holistic understanding of the relationships between people, processes, and technology. In the context of batch processing in public institutions, this means that centralized process scheduling requires both technical and organizational adaptations—from the standardization of procedures to the training of operators.

Another important theoretical foundation is contingency theory, which asserts that there is no universally optimal organizational design; rather, the effectiveness of specific practices depends on the given context. Lawrence and Lorsch (1967) demonstrated that the success of organizational structures arises from their adaptation to environmental complexity, while Donaldson (2001) expanded this perspective within a broader understanding of organizational contingencies. In the field of operations management, Sousa and Voss (2008) provided empirical evidence that structured planning yields the greatest benefits under conditions of high interdependence and uncertainty. More recently, scholars such as Zheng (2025) have developed adaptive scheduling models that respond dynamically to unexpected disruptions in real time. For public institutions, where the environment is often unpredictable and systemic constraints are significant, such an approach is particularly relevant.

The next body of theory derives from process optimization and the Theory of Constraints (TOC). Goldratt and Cox (1984) demonstrated that an organization's effectiveness is determined primarily by its ability to identify and manage bottlenecks. Within the field of workflow management, Van der Aalst and Van Hee (2004) developed models for analysing and optimizing task dependencies, while Méndez et al. (2006) showed through a comprehensive review of methods that systematic scheduling of batch processes significantly improves throughput and reduces processing times. More recent research (e.g., Feng et al., 2024; Zheng, 2025) has expanded this approach by incorporating multi-objective goals such as energy efficiency and cost reduction, indicating that process optimization extends beyond purely technical improvements to encompass broader aspects of sustainable organizational performance.

An additional perspective is offered by Structuration Theory, developed by Giddens (1984). This theory is based on the assumption that structures shape individual behaviour, while at the same time individuals reproduce and transform those structures through their actions. Orlikowski (1992) demonstrated that technology in organizations does not exist as a neutral tool but instead acquires meaning through its enactment in practice. In public institutions, this implies that the introduction of centralized batch scheduling represents not only a technical innovation but also a reorganization of work practices that evolves through the interaction of employees with the system.

Equally important is the Social Construction of Technology (SCOT) theory, developed by Pinch and Bijker (1987). According to their view, technology does not operate in a deterministic manner; rather, its meaning and function emerge through the interpretations and negotiations of different social groups. In the case of public institutions, this suggests that policymakers, information system developers, administrators, and end-users may each perceive and evaluate centralized scheduling in different ways, which in turn influences its actual effectiveness. This perspective explains why the introduction of centralized scheduling may successfully reduce processing times but does not necessarily increase process reliability, as this outcome also depends on the social acceptance and integration of the technology into everyday work practices.

4. Hypotheses

In this study, the notion of "planning the execution of batch processing" refers to the set of activities and mechanisms introduced with the centralized batch management system. Concretely, this includes: defining and maintaining a centralized schedule of batch jobs; configuring temporal conditions (time windows, execution frequency) and logical dependencies (precedence constraints, prerequisite checks); linking batch jobs to standardized work orders and documentation; and enabling operators to monitor and adjust execution via a dedicated interface. Planning, as operationalized in our analysis, is therefore not an abstract concept but a socio-technical configuration in which temporal rules, dependency management, and organizational procedures are explicitly encoded into the system.

The objective of this research is to empirically test three hypotheses related to the impact of processing duration on success rates, the influence of planning on duration, and the influence of planning on the success of batch processes. Through quantitative analysis of real processing log data from the period 2018–2023, the article aims to demonstrate opportunities for process optimization in data-intensive environments and offer practical recommendations for improving information infrastructure within the public sector.

H1: The Duration of Batch Processing Affects Its Success.

Hypothesis H1 is based on several theories and empirical observations confirming that processing duration is a key indicator of process stability, resource availability, and execution success. Authors such as McHugh, Shiff, and Gaur emphasize that prolonged processing often reflects structural or systemic deficiencies that reduce system reliability.

According to McHugh (2024), processing time is a direct indicator of system efficiency, as longer durations often lead to bottlenecks and system overload. While each process does not require the same execution time, unusually long durations frequently signal failures in the processing chain, especially when involving complex technical dependencies or high system loads.

Shiff (2024) emphasizes that in sequentially dependent batch processes, timing synchronization is a critical factor. If one task takes too long, it disrupts the entire schedule, increasing the risk of failure for subsequent tasks, particularly those that are time-sensitive or depend on data from earlier stages. Such delays are not merely postponements; they can lead to invalid calculations or missing data, ultimately resulting in failed processing.

Mastering Batch Processing (2024) highlights that in systems with complex technical or logical dependencies, processes often execute more slowly due to waiting for resources or confirmations. When these dependencies are not properly aligned, task durations increase, which can cause downstream process failures. This not only extends the total system time but also raises the likelihood of cascading errors.

Gaur (2024) stresses the importance of log data analysis for identifying patterns between duration and failure. In practice, processes that exceeded the average execution time by more than 25% had nearly twice the likelihood of failure, indicating that duration is a reliable risk indicator. This directly supports the theoretical foundation of Hypothesis H1.

H2: Planning the Execution of Batch Processing Reduces Processing Time

Hypothesis H2 is based on the assumption that structured, pre-defined planning of batch processes enables better resource distribution, eliminates unnecessary delays, and ensures an optimal execution sequence. As a result, it shortens the overall processing time. This link between planning and processing duration has theoretical roots in workflow management, system optimization, and job scheduling theory.

Shorten (2024) emphasizes that effective scheduling is a key factor in reducing processing time in data-intensive environments. In systems lacking clear temporal and sequential planning, delays often occur due to resource contention, task conflicts, and duplicated procedures—all of which directly extend execution time. In contrast, dynamic planning and intelligent task orchestration allow for adjusting task order based on priorities, current system load, and historical performance.

Gaur (2024) highlights that systems like Apache Airflow and IBM Tivoli Workload Scheduler have demonstrably reduced process durations by enabling: dependency recognition, automatic scheduling based on optimal sequencing, and parallel execution where logical constraints allow. This shows that planning is not merely an administrative task but an active optimization mechanism that directly influences the temporal efficiency of processes.

nOps (2024) further points out that one of the most common causes of extended processing times is an unoptimized task sequence. When a system cannot predict resource availability or identify independent tasks, it often defaults to unnecessary serial execution, where parallelism would be possible. By introducing planning mechanisms that consider resource availability, time windows, and critical paths, bottlenecks can be eliminated, leading to significantly shorter execution times.

H3: Planning the Execution of Batch Processing Affects Processing Success

Hypothesis H3 is based on the assumption that structured and deliberate planning of batch executions positively influences their success by enabling better system preparation, conflict resolution, error reduction, and improved availability of data and resources. Numerous studies support the idea that process organization significantly contributes to the reliability and quality of processing outcomes.

According to Johnson (2024), planning in data processing systems is not merely a temporal function but also a qualitative mechanism for ensuring stability and predictability. When the execution order is clearly defined, dependencies are resolved in advance, and the system is aware of the time windows during which tasks are executed, errors related to missing input data, resource access conflicts, or human mistakes during manual triggering are greatly reduced. Precise planning thus fosters higher process discipline, increasing the likelihood of successful completion.

Talend (2024) states that planning also introduces procedural standardization, meaning that processes are executed under consistent rules, at predictable times, and with the appropriate resource allocation. This consistency significantly lowers the risk of errors stemming from improvisation or undocumented interventions. From a quality management perspective, planning is therefore a prerequisite for success.

Nicole (2024) emphasizes that in systems where processing is initiated without planning, the probability of unexpected events, such as interruptions, unresolved dependencies, or system overload, increases. These conditions heighten the risk of failure. In contrast, planning introduces predictability, enabling systems to reserve resources appropriately and resolve potential conflicts, thus contributing to higher processing success rates.

Maronde (2024) notes that the success of processing is not solely dependent on planning, but also on factors such as operator competence, input data quality, and infrastructure reliability. Nevertheless, planning helps reduce errors caused by organizational deficiencies, functioning as a stabilizing factor even in complex environments.

While H1–H3 capture baseline expectations derived from the literature regarding duration, planning, and success, the empirical rejection of H1 and H3 suggests that other factors may be more decisive for processing success. Instead of formulating additional a priori hypotheses, we therefore complement our analysis with exploratory research questions for future work, such as:

- Which non-temporal factors (e.g., data quality, infrastructure stability, operator interventions) best explain variations in processing success?

- How do different configurations of planning (e.g., stricter dependency rules vs. more flexible scheduling) affect trade-offs between efficiency and robustness?

These questions extend the current hypothesis set and provide a roadmap for subsequent studies.

Although planning did not significantly increase processing success, the results highlight the importance of operator oversight and structured work orders. Unsuccessful cases were more frequently associated with inadequate input data, manual handling errors, or unrecorded dependencies—factors unrelated to system scheduling. This indicates that technological improvements must be complemented by organizational adjustments, including clearer documentation, operator training, and standardized work order procedures, consistent with socio-technical systems theory.

5. Methodology

Methodologically, the research was structured into four main phases. In the first phase, the proprietary centralized batch management system was designed and implemented, encapsulating the planning mechanisms described in Section 4. In the second phase, log data from 2018 to 2023 were extracted, cleaned, and segmented into pre-implementation and post-implementation periods, enabling a natural comparison between “non-planned” and “planned” execution regimes. The third phase consisted of descriptive analysis of key variables, processing duration, success rates, and processing types, providing a quantitative overview of the institutional workload. In the fourth phase, we conducted inferential statistical tests to evaluate Hypotheses H1–H3. In this way, each hypothesis is directly linked to a specific operationalization in the data and to a clearly defined analytical step.

The development phase of the research focused on the implementation of centralized batch process management. For this purpose, a dedicated user interface was developed in the C# programming language, enabling the management, monitoring, and triggering of batch processes. The system was tightly integrated with the existing Oracle database, allowing real-time access to metadata and processing logs.

The main functionalities of the developed system include:

- automatic dependency checks between processes,
- triggering based on temporal and logical conditions,
- visualization of schedules and tasks, and
- alerting in the event of processing failures.

To evaluate system performance, a database of batch processing logs from 2018 to 2023 was used. The dataset included the following elements:

- start and end timestamps of each process,
- process identifier,
- execution outcome (successful/unsuccessful),
- processing type (manual/automatic),
- associated work order.

Before analysis, the data had to be cleaned, removing missing, inconsistent, and duplicate records. The data were then segmented into two periods: before and after the implementation of the new system. The system was implemented in early 2020.

The proprietary batch management system does not implement adaptive or self-optimizing scheduling algorithms. Instead, it operates as a centralized, rule-based system that enforces predefined temporal schedules, logical dependencies, and execution conditions. Its primary function is to replace fragmented manual coordination with a unified control mechanism, ensuring consistent execution order, explicit dependency checks, and improved visibility for operators. The efficiency gains observed in the analysis therefore stem from reduced manual intervention, clearer sequencing of tasks, and the elimination of ad hoc execution practices rather than from dynamic optimization or predictive decision-making.

This study has several limitations that must be considered when interpreting the results. The research was conducted within a single public institution, which means that the findings cannot be directly generalized to all public organizations. The dataset used is based primarily on technical process logs, which provide precise measurements of efficiency but do not capture softer factors such as user satisfaction, employee competencies, or organizational culture. Methodologically, the study relies mainly on a quantitative approach, which limits the depth of contextual understanding that could have been enriched through qualitative methods. Furthermore, the analysis covers a specific time period, implying that subsequent technological or organizational changes may affect the replicability of the results. The analysis covers the period from 2018 to 2023, with 2018–2019 representing the pre-implementation period and 2020–2023 the post-implementation period.

6. Results

Log Data Analysis

The log data serve as a critical component for monitoring and analyzing batch processes at the institution. This system records all essential information, including process start and end times, execution outcomes, and processing results. These logs provide not only a clear overview of process performance but also valuable insights for identifying errors and enhancing procedures in the future.

The analysis of the institution's log data also reveals a significant increase in the number of batch processes, particularly those executed on a scheduled basis. The data clearly show a year-over-year growth in the volume of automatically executed processes, while the number of manual processes has remained relatively stable.

Year	Scheduled Processing	Manual Processing
2018	41.562	18.570
2019	48.420	19.250
2020	53.516	17.557
2021	135.326	18.130
2022	195.464	16.978
2023	312.054	17.689

Table 1. Growth in scheduled and manual processing (2018-2023).

Table 1 presents the number of scheduled and manual batch processes from 2018 to 2023. The analysis of the data shows a clear and significant increase in scheduled processing over the observed period, while manual processing remained relatively stable with minor fluctuations. Between 2018 and 2019, scheduled processing increased by approximately 16.49%, whereas manual processing rose by only 3.66%. In 2020, scheduled processing continued to grow by 10.52%, while manual processing decreased by 8.79%. The most significant increase in scheduled processing occurred between 2020 and 2021, with a 152.82% rise in the number of executions. In contrast, manual processing increased by just 3.27%. In 2022, scheduled processing grew again by 44.43%, while manual processes declined by 6.36%. From 2022 to 2023, scheduled processes surged once more by 59.66%, while manual processes showed a slight increase of 4.19%. These figures clearly indicate a continuous upward trend in scheduled processing, particularly after 2020, reflecting greater automation and centralization of batch execution. In contrast, manual processing remained relatively stable, with only modest changes over time.

Analysis of Conditions and Dependencies

Conditions and dependencies are critical elements in the execution of batch processes, as they directly influence the correctness, sequence, and success of individual tasks. Every batch process depends on specific conditions that must be met to ensure a proper and orderly execution. These conditions may include time constraints, the status of preceding processes, the availability of input data, or predefined parameters that must be fulfilled before the process can start.

In the institution's system, these conditions are explicitly recorded in a dedicated column, where they are clearly defined for each individual process. This field contains information about the prerequisites that must be met prior to execution. Often, these dependencies relate to previously completed processes or the status of system resources—such as file availability, access to data sources, or other processing requirements.

For manual processes, where an operator oversees execution, these conditions are especially critical. The operator must verify that all prerequisites are fulfilled before initiating a task. For instance, if a process is dependent on the outcome of another task, it is essential that the operator confirms the successful completion of the preceding process. Failure to meet these conditions can result in errors or incorrect outputs.

Processing ID	Conditions	Pre-Processing	Pre-Processing Performance	Status
1001	Conditions A	1000	Successful	OK
1002	Conditions B	1001	Failed	STOP

Table 2. Conditions and dependencies between processes.

Table 2 illustrates how individual batch processes are interconnected and how the success of preceding processes affects the overall workflow. Each process is subject to certain conditions and is dependent on the successful completion of a prior process. If the previous process is successful, the next one proceeds; otherwise, the sequence is halted.

Table Columns:

- Processing ID: Unique identifier for each batch process.
- Conditions: Specific requirements that must be fulfilled before execution.
- Pre-Processing: The process that must successfully complete beforehand.
- Pre-Processing Performance: Indicates whether the prior process succeeded.
- Status: Whether the current process can proceed ("OK") or is blocked ("STOP").

In automated processing, these conditions are verified through system rules and schedules. Automated condition checking ensures that processes are only triggered when all necessary criteria are met. The system evaluates: whether previous tasks have completed successfully, if all required input data is available, and whether any time constraints are satisfied.

Process Dependencies play a vital role because many batch jobs are linked in chains or sequences. A successful process execution is often a prerequisite for the next one to begin. Dependency analysis enables identification of critical control points, where potential failures or delays could cascade through the entire batch sequence. This makes understanding dependencies essential for effective planning and optimization of task execution.

At the institution, dependencies are either temporal or logical. Temporal dependencies relate to when tasks must run, at specific intervals or in a defined order. It is crucial that scheduling prevents resource conflicts (e.g., database or network contention). Logical dependencies are tied to the outcome of previous processes. A task may only proceed once the prior process has completed successfully and the necessary data is prepared.

Managing these dependencies and conditions is essential for the smooth operation of the batch processing system. Analysing dependencies and preconditions allows the system to: accurately determine execution order, prevent errors caused by improper sequencing, and reduce the risk of delays and system faults.

By systematically analysing conditions and dependencies, organizations can significantly improve the management of batch processing, leading to greater reliability, efficiency, and overall system performance.

Before proceeding with hypothesis testing, a descriptive analysis of the data must be conducted to gain a foundational understanding of the sample characteristics. This step involves summarizing key statistical indicators that provide insights into the structure and variability of the data.

The following metrics will be calculated:

- Means (Averages) – to determine the central tendency of processing times, success rates, and other key variables.
- Standard Deviations – to assess the variability and dispersion within the dataset.
- Distributions – to observe the frequency and patterns of specific values (e.g., successful vs. unsuccessful executions, scheduled vs. manual processing).

This initial overview will help identify trends, outliers, and potential anomalies, and will provide the necessary context for a more accurate and meaningful hypothesis testing process.

	Duration of processing execution in minutes on a work order
Number (of work orders)	1522
Mean	161,64717
Standard deviation	80,21125
Minimum	23
Maxsimum	473
25 %	80
50 %	184
75 %	221

Table 3. Descriptive statistics on processing duration on work orders.

Table 3 shows data on the duration of batch processing in minutes per work order, with a total of 1,522 work orders analysed. The average processing duration is 161.65 minutes, indicating that most processes take slightly more than two and a half hours. The standard deviation is 80.21 minutes, which means that the duration of processing can vary significantly between individual tasks. The shortest process lasted 23 minutes, while the longest took 473 minutes. Table 3 also presents the quartiles of processing duration. One quarter of the processes lasted 80 minutes or less, half of the processes were completed in 184 minutes or less, and one quarter of the processes lasted more than 221 minutes.

	Work order performance
Number (of work orders)	1522
Number of successful	1406
Number of unsuccessful	116
Minimum	0 - Successful
Maximum	1 – Unsuccessful

Table 4. Descriptive statistics on the success of work order processing.

Table 4 shows the success rate of work orders, with a total of 1,522 orders analysed. Out of these, 1,406 were successfully executed, representing approximately 92.38% of all orders, while 116 orders failed, accounting for about 7.62%. The minimum success value is 0, meaning that in a certain period all orders were unsuccessful, while the maximum value is 1, meaning that in a certain period all orders were successfully executed.

	Duration of processing execution in minutes on a work order
Number (of work orders)	976
Mean	14,46926
Standard deviation	22,09812
Minimum	0
Maxsimum	148
25 %	3
50 %	6
75 %	15

Table 5. Descriptive statistics on the duration of selected treatments from 2018 to 2023.

Table 5 presents a descriptive analysis of two key variables related to the selected processes: processing duration and processing success. The analysis is based on a sample of 976 selected processes. The average processing duration is 14.47 minutes, with a standard deviation of 22.10 minutes, indicating high variability in the duration of individual processes. The shortest process lasted 60 seconds, while the longest process lasted 148 minutes. One quarter of all processes lasted less than 3 minutes, half of the processes were completed in less than 6 minutes, and one quarter lasted more than 15 minutes. The results show that the majority of processes were successful and that the duration of individual executions varied considerably, with the average time being slightly under 15 minutes.

	Treatment performance	Treatment performance selected
	Frequency	Percentage
Number of treatments	976	100 %
Number of successful	975	99,9 %
Number of unsuccessful	1	0,01 %

Table 6. Descriptive statistics on the success of selected treatments from 2018 to 2023.

The table 6 presents the success rate of the selected processes, with a total of 976 processes analysed. Out of these, 975 were successful, which means that the vast majority of processes were completed successfully. Only one process failed, indicating a very high success rate.

H1: The duration of batch processing affects its success.
In this analysis, we tested Hypothesis H1, which claims that the execution time of batch processes influences their success. By comparing the time frames of process execution with their outcomes, we examined whether there is a statistically significant relationship between processing duration and success.

Performance	Mean (in minutes)	Standard deviation	N	ANOVA F- statistics	Sig.
Successful (0)	162,21692	79,32931	1406	0,93073	0,33482
Unsuccessful (1)	154,74137	90,31482	116		

Table 7. Comparison of duration of successful and unsuccessful treatments (2018-2023).

Table 7 presents a comparison of the average duration between successful and unsuccessful processes. The average duration of successful processes is 162.22 minutes with a standard deviation of 79.33 minutes, while unsuccessful processes last on average 154.74 minutes, with a standard deviation of 90.31 minutes. The number of successful processes was 1,406, while there were 116 unsuccessful processes. The results of the ANOVA analysis show an F-statistic of 0.93073 and a p-value of 0.33482, which indicates that there is no statistically significant difference between the duration of successful and unsuccessful processes. Based on the statistical analysis, we reject H1, as the duration of batch processes does not affect their success.

The rejection of Hypotheses 1 processing time is one of the key predictors of success in data-intensive systems; Gaur (2024), who claims that longer durations are often associated with a higher number of errors due to bottlenecks; and Shiff (2024), who highlights the cascading effect of delays leading to errors in dependent processes.

In the specific case of the analysed public institution, we may conclude that the quality of process execution was likely more dependent on other factors, such as infrastructure stability, input data quality, or operator supervision, rather than on the duration of the tasks themselves.

H2: Planning the Execution of Batch Processing Reduces Processing Time.
Hypothesis H2 is based on the assumption that structured planning of batch processing has a significant impact on time efficiency. This thesis is supported by several authors, such as Shorten (2024), Gaur (2024), and nOps (2024), who emphasize that well-scheduled processing tasks enable better utilization of resources, prevent waiting times, and eliminate unnecessary temporal overlaps within complex data workflows.

Period	Mean (in minutes)	Standard deviation	N	ANOVA F- statistics	Sig.
Before 2020	207,85317	60,75913	504	299,04597	0,000
After 2020	138,77111	78,83241	1018		

Table 8. Comparison of processing times before and after 2020.

Processes were grouped into two periods: before 2020 and after 2020. The transition year 2020 is included in the post-implementation dataset, as proprietary centralized batch management system was introduced early that year. The table presents a comparison of the average duration of batch processes before 2020 and after the introduction of improved planning. Before 2020, the average duration of processes was 207.85 minutes with a standard deviation of 60.76 minutes, indicating longer but more consistent execution times. After 2020, the average duration decreased to 138.77 minutes, reflecting a reduction of approximately 69 minutes per process. The number of processes before 2020 was 504, while after 2020 it increased to 1,018. The results of the ANOVA analysis show an F-statistic of 299.05 and a p-value less than 0.5, which confirms that Hypothesis H2 can be accepted. We can conclude that planning of batch processes has a statistically significant effect on reducing execution time. After the introduction of advanced planning, the average execution time decreased by 69 minutes, confirming that careful process planning plays a crucial role in improving efficiency and accelerating execution.

These results are consistent with the findings of: Shorten (2024), who notes that real-time intelligent planning leads to shorter execution times in data systems. nOps (2024), who highlights that more than 40% of processing delays are caused by non-optimal task sequencing. Gaur (2024), who confirms that dynamic dependency management increases system throughput.

In this research, the results confirm that advanced planning has a key impact on reducing process duration, as it enables better scheduling control, timely activation of dependent tasks, and optimal coordination in highly complex systems. This is particularly important in public institutions, where process stability and time synchronization directly affect the quality of services.

H3: Planning the Execution of Batch Processing Affects Processing Success.
In this analysis, we tested Hypothesis H3, which states that planning of batch processes affects their success. By comparing process success before and after the implementation of planned processes, we analysed whether the introduced changes led to fewer errors, improved timeliness of execution, and greater user satisfaction.

Period	The proportion of successful
Before 2020	92,06349
After 2020	92,53438

Table 9. Comparison of successful processing before and after implementation of planning.

The table presents a comparison of the share of successful processes before 2020 and after the implementation of improved planning. Before 2020, the share of successful processes was 92.06%. After 2020, this share slightly increased to 92.53%. Although the difference is small, the data indicate a slight improvement in success following the implementation of planned processes. Based on these results and the minimal difference in success rates, it can be concluded that planning did not have a significant impact on process success. On the basis of statistical analyses, Hypothesis H3 is rejected, meaning that the planning of batch processing execution does not significantly influence process success.

Taken together, the results for H2 and H3 suggest an important asymmetry in the effects of planning as operationalized in this study. On the one hand, planning, implemented through centralized scheduling, explicit temporal rules, and dependency management, has a strong and statistically significant effect on execution time (H2), confirming that such mechanisms are effective levers for improving temporal efficiency. On the other hand, the same planning configuration has only a negligible effect on success rates (H3). This indicates that "planning the execution", as we define it here, primarily addresses when and in which order tasks are executed, but does not sufficiently intervene in other determinants of success, such as input data quality, infrastructure reliability, and exceptional event handling. In other words, the planning regime is effective in smoothing the temporal dimension of processing but only weakly influences the deeper socio-technical sources of reliability.

This finding suggests that the effect of planning on success is less pronounced than its effect on duration, which aligns with the conclusions of Johnson (2024), who argues that planning influences efficiency more strongly than it does execution quality, except in highly complex or critical workflows. Similarly, Maronde (2024) points out that organizational changes, such as the introduction of planning, may have limited impact on success if the main causes of failures lie elsewhere, for instance, in data quality, infrastructure robustness, or unpredictable system outages.

The results of this study confirm that centralized and structured scheduling of batch processes significantly reduces execution time and thereby increases the efficiency of information systems in a public institution. This finding advances existing knowledge by providing empirical evidence consistent with process optimization and workflow management theories, which emphasize the importance of eliminating bottlenecks and improving task sequencing. However, the study also shows that scheduling has only a limited impact on overall process success, indicating that factors such as infrastructure robustness, data quality, and operator competence remain equally important. This nuance extends previous research, which often assumes a stronger link between planning and reliability, by demonstrating that in the context of public institutions, efficiency gains do not automatically translate into higher levels of success.

Although the finding that centralized scheduling reduces execution time may appear intuitive, its empirical confirmation in a public-sector environment is non-trivial. Many studies implicitly assume that automation and planning also improve reliability; however, the rejection of H1 and H3 provides empirical evidence against this assumption. The results demonstrate that, in complex institutional settings, efficiency gains achieved through centralized scheduling do not automatically translate into higher success rates. This negative finding is theoretically relevant, as it challenges simplified narratives of automation-driven reliability and underscores the need to distinguish between temporal efficiency and execution robustness.

By critically reflecting on these findings, the study contributes to the broader theoretical debate in the literature on organizations and information systems. Specifically, it demonstrates that socio-technical alignment is essential: improvements in efficiency require not only technological solutions but also organizational adaptation, training, and standardized procedures. The rejection of Hypotheses 1 and 3 highlights the need to revisit assumptions within contingency theory and structuration theory, emphasizing that process success in highly complex and interdependent environments depends on more than scheduling mechanisms alone. Concretely, our findings suggest at least three directions for such refinement. First, future contingency-theoretic work should explicitly model infrastructure robustness, data quality, and regulatory constraints as contextual variables that moderate the impact of planning mechanisms, using comparative designs across institutions with different technical and organizational profiles. Second, structuration-inspired studies could investigate how operators appropriate or circumvent centralized scheduling tools in practice, for instance through qualitative case studies or mixed-method designs that combine log analysis with interviews and observations. Third, empirical criteria for evaluating success should be broadened beyond binary execution outcomes to include measures such as recovery time after failures, stability across peak-load periods, and the effort required from human operators. These approaches would allow a more fine-grained assessment of when and how centralized scheduling creates value, and where complementary interventions are necessary.

Beyond the specific institutional case, the results have broader implications for other public-sector and organizational contexts facing similar digitalization challenges. Many government agencies and large organizations operate with legacy systems characterized by high interdependencies and limited resources. The evidence presented here suggests that centralized scheduling can be an effective intervention to reduce delays and increase efficiency in such contexts, even if it does not fully resolve issues of reliability. Thus, the study provides practical insights for policymakers and practitioners while also offering a framework that can be generalized to a wide range of organizations seeking to modernize their information infrastructures through structured process management.

The results of the analysis clearly show that non-temporal factors play a decisive role in explaining variations in processing success. Failures most commonly arise from issues such as incomplete or inconsistent input data, infrastructure instability, database locks, and fluctuations in system load, factors that are independent of execution timing or scheduling quality. Operator-related aspects, including manual overrides, mis-sequenced task initiation, or deviations from documented procedures, also significantly influence outcomes. Furthermore, processes embedded in long or fragile dependency chains are particularly vulnerable, as any upstream interruption propagates across subsequent tasks regardless of the underlying planning regime. These findings demonstrate that the determinants of processing success lie primarily in data quality, infrastructure robustness, operator competence, and the structural characteristics of dependency networks, which explains why planning mechanisms substantially improved temporal efficiency (H2) but did not meaningfully affect overall success rates (H3).

The analysis also highlights that different configurations of planning introduce varying trade-offs between efficiency and robustness. Stricter planning configurations, such as rigid dependency enforcement, comprehensive validation of prerequisites, and narrow execution windows, enhance robustness by restricting execution to conditions that minimize the risk of failure. However, these constraints reduce flexibility, limit parallelism, and may increase overall processing time. In contrast, more flexible configurations allow broader time windows, dynamic triggering, and greater concurrency, which improves throughput but increases exposure to resource contention, synchronization problems, and reduced operator oversight. A hybrid approach, in which critical or failure-sensitive processes follow stringent rules while routine tasks operate under more flexible conditions, offers a balanced solution. This indicates that the design of planning mechanisms matters as much as their presence: planning must be tailored to the technical, organizational, and operational characteristics of the environment if institutions are to achieve both efficiency gains and improved reliability.

7. Suggestions for Improvements

The proposed recommendations are directly derived from the empirical findings and the theoretical framework adopted in this study. Since the analysis shows that failures are primarily associated with non-temporal factors, such as data quality, infrastructure instability, and operator interventions, rather than with scheduling mechanisms, the suggested improvements deliberately focus on enhancing informational transparency, analytical support, and organizational coordination rather than on further tightening execution schedules.

Based on the analysis conducted and the review of current processes, the following improvements are recommended for the Institution:

1. **Enhancement of Log Records**:
   Currently, log records mainly capture timestamps (start and end of processing) and execution success. While this provides basic monitoring, it would be beneficial to extend records with more detailed information. Standardized and richer logs that include all necessary process parameters would allow deeper insights and advanced analytics, such as machine learning and artificial intelligence. This would make the system more modern and prepared for future technological trends, while also supporting further process optimization. In particular, logs should include

information on resource usage (memory), error codes, dependency identifiers, data input validation results, operator actions, queue wait times, and system load at execution time. These fields would enable predictive analytics and support the identification of structural failure patterns.

2. **Standardization of Records**:
To ensure more effective use of data in the future, log records should be standardized. This would make analysis consistent regardless of the type or scope of processing. Standardization would simplify historical process analysis, help predict potential issues, and support more accurate planning of future processes.
3. **Development of Analytical Tools**:
To fully leverage past process data, tools should be developed to predict process success and duration based on historical records. This would allow more precise planning and reduce the likelihood of future errors. To fully leverage past process data, the institution should develop analytical tools that provide early-warning signals and decision support based on historical patterns. For example, dashboards could visualize distributions of processing times and failure rates by job type, time of day, or dependency chain, enabling operators to detect emerging bottlenecks. Statistical or machine-learning models could estimate the probability of failure for an upcoming execution based on features such as past duration, recent infrastructure incidents, or the length of the dependency chain. Such tools would not replace operators but would support them with concrete risk scores and alerts, allowing them to prioritize monitoring and preventive interventions for the most vulnerable processes.
4. **Use of Artificial Intelligence and Machine Learning**:
Artificial intelligence and machine learning can further enhance batch management by learning complex patterns that are difficult to capture with simple rules. For instance, anomaly detection algorithms could automatically flag executions whose duration or resource usage deviates significantly from historical norms, even if they have not yet failed. Sequence models could identify combinations of jobs and temporal conditions that are systematically associated with later disruptions. Recommendation systems could propose more robust scheduling configurations (e.g., alternative time windows or parallelization options) based on past successes. To make such approaches feasible, the institution must first ensure standardized and high-quality logs, as outlined in points 1 and 2, and then iteratively validate AI-generated insights with operators before automating any actions.
5. **Centralized Management and Planning**:
Analysis results have shown that centralized management and careful planning significantly reduce operational delays and improve process performance. With an even more systematic approach to planning, processes could be further optimized, reducing the risk of failures.

Log records from past processes have great potential for improving decision-making in proprietary centralized batch management systems. However, technical and content upgrades are necessary to unlock this potential. Unified and more detailed records would enable the use of advanced technologies, such as AI and machine learning, to optimize processes, making the system more aligned with modern trends and better equipped to predict and resolve issues in a timely manner.

Nevertheless, challenges remain, such as incomplete or inconsistent data in the current logs, which limits forecasting and optimization capabilities. For effective use of this data, consistency and accuracy must be ensured, along with robust security and protection mechanisms. Moving forward, investments should focus on improving log quality and developing tools that enable efficient forecasting and process optimization.

8. Conclusion

The research has shown that the introduction of advanced, centralized management significantly contributed to the reduction of batch processing duration in the analysed public institution. After the implementation of the centralized planning system, processing times decreased noticeably, confirming that careful planning and centralized process management have a positive impact on operational efficiency. At the same time, the results did not indicate a statistically significant effect on overall success rates, as the proportion of failed processes remained almost unchanged before and after the implementation. This asymmetry suggests that temporal optimization and reliability are shaped by partly different mechanisms.

This study does not claim to offer a breakthrough in automated batch scheduling or dynamic optimization. Instead, its contribution lies in providing a realistic and empirically grounded assessment of centralized batch management in a public-sector context. The findings show that while centralized, rule-based planning is an effective tool for reducing execution time, it has limited influence on processing success, which depends on broader socio-technical conditions. By explicitly documenting these limitations, the paper contributes to a more balanced understanding of digital transformation in the public sector and cautions against overstating the capabilities of automation without corresponding organizational and infrastructural investments.

From a practical standpoint, the findings offer several benefits for stakeholders. For IT managers and system architects, the study provides quantitative evidence that investments in centralized scheduling and dependency management can yield substantial time savings, even in environments with legacy systems and limited resources. For operators, the results highlight where planning helps, by reducing manual interventions and clarifying execution sequences, and where it does not, namely in preventing failures caused by data quality issues or infrastructure incidents. For policymakers and organizational leaders, the study clarifies that digitalization efforts focused solely on automation and scheduling will improve efficiency but will not, on their own, guarantee higher reliability or service quality.

The results reported in this paper can be developed further in several directions. First, future work should extend the analysis beyond a single institution and compare public organizations with different technical infrastructures and governance models; to test how contextual factors moderate the effects observed here. Second, incorporating qualitative data, such as interviews with operators, administrators, and users, would allow a deeper understanding of how centralized scheduling is appropriated in practice and how informal workarounds influence outcomes. Third, the proposed improvements in logging and analytical tools could be implemented and evaluated in a pilot setting, for example by testing whether predictive models and anomaly detection reduce the number or impact of failures over time.

In a broader perspective, the study contributes to the ongoing discussion on digital transformation in the public sector. It demonstrates that centralized scheduling is a powerful instrument for improving efficiency, but not a universal remedy for all reliability problems. Achieving sustainably reliable batch processing requires a combination of technological enhancements (centralized management, better logging, analytical tools), organizational measures (clearer work orders, training, process standardization), and human resource development. Only by jointly addressing these technical and organizational dimensions can public institutions fully realize the potential of advanced batch management and move towards more resilient, data-intensive service provision.

9. References

Chia, L. (2024). Batch Processing Analysis in Large Institutions: SWOT Perspective. Journal of Information Systems, 33(2), 123–135.

Donaldson, L. (2001). The contingency theory of organizations. Sage Publications.

Feng, Y., Zhang, H., & Sun, X. (2024). Energy-efficient batch scheduling with makespan and electricity cost minimization. Mathematics, 12(3), 376.

Gaur, V. (2024). Efficient Workflow Orchestration in Public Systems. Systems Engineering Review, 40(1), 45–63.

Giddens, A. (1984). The constitution of society: Outline of the theory of structuration. University of California Press.

Goldratt, E. M., & Cox, J. (1984). The goal: A process of ongoing improvement. North River Press.

IBM. (2024). Modern Batch Processing Systems and Job Control Language. IBM Knowledge Center.

https://www.ibm.com/docs

Johnson, T. (2024). Cloud-Based Optimization of Batch Workflows. Proceedings of the IEEE CloudTech Conference, 2024.

Khalfe, R. (2024). Transaction Data Flows and Batch Dependencies in Financial Institutions. FinTech Data Journal, 29(3), 188–201.

Lawrence, P. R., & Lorsch, J. W. (1967). Organization and environment: Managing differentiation and integration. Harvard Business School Press.

Maronde, P. (2024). Organizational Change and IT System Implementation. Public Sector Informatics Review, 15(1), 59–74.

Mastering Batch Processing. (2024). Principles of Batch Job Scheduling and Execution.

https://www.masteringbatch.org

McHugh, D. (2024). Automation and Resource Optimization in Data-Intensive Institutions. Information Management Quarterly, 12(4), 212–230.

Méndez, C. A., Cerdá, J., Grossmann, I. E., Harjunkoski, I., & Fahl, M. (2006). State-of-the-art review of optimization methods for short-term scheduling of batch processes. Computers & Chemical Engineering, 30(6–7), 913–946.

Mumford, E. (2006). The story of socio-technical design: Reflections on its successes, failures and potential. Information Systems Journal, 16(4), 317–342.

Nicole, B. (2024). Data Integrity and Backup Strategies for Critical Systems. Journal of Data Safety, 6(2), 88–97.

nOps. (2024). Error Resilience and Dependency Mapping in Batch Workflows. White Paper Series.

https://www.nops.io/resources

Orlikowski, W. J. (1992). The duality of technology: Rethinking the concept of technology in organizations. Organization Science, 3(3), 398–427.

Pinch, T. J., & Bijker, W. E. (1987). The social construction of facts and artifacts: Or how the sociology of science and the sociology of technology might benefit each other. V W. E. Bijker, T. P. Hughes & T. J. Pinch (ur.), The social construction of technological systems (17–50). MIT Press.

Roth, S. (2023). Sociotechnical perspectives in Industry 4.0: Aligning technology and organizational practices. Sustainability, 15(23), 16438.

Shiff, M. (2024). Dependency Management in Job Scheduling Systems. ACM Transactions on Information Systems, 41(1), 1–29.

Shorten, C. (2024). Predictive Batch Processing using Real-Time Analytics. DataOps Engineering Reports, 18(2), 98–115.

Sousa, R., & Voss, C. A. (2008). Contingency research in operations management practices. Journal of Operations Management, 26(6), 697–713.

Talend. (2024). Data Pipeline Orchestration for Hybrid Environments.

https://www.talend.com/resources

Thomas, E. (2024). Foundations of Batch Processing and System Optimization. IT Operations Handbook, Vol. 2, 70–91.

Trist, E. L., & Bamforth, K. W. (1951). Some social and psychological consequences of the longwall method of coal-getting. Human Relations, 4(1), 3–38.

Van der Aalst, W. M. P., & Van Hee, K. M. (2004). Workflow management: Models, methods, and systems. MIT Press.

Zheng, S. (2025). Bi-objective scheduling of non-identical batch machines considering energy efficiency. Annals of Operations Research, 346(3), 145–167.