Proceedings of the 2025 Winter Simulation Conference
E. Azar, A. Djanatliev, A. Harper, C. Kogler, V. Ramamohan, A. Anagnostou, and S. J. E. Taylor, eds.

AUTOMATED BUSINESS PROCESS SIMULATION STUDIES: WHERE DO HUMANS FIT IN?

Samira Khraiwesh[1], and Luise Pufahl[1]

[1]Dept. of Computer Science, Technical University of Munich,
Heilbronn, GERMANY

ABSTRACT

Business Process Simulation (BPS) is crucial for enhancing organizational efficiency and decision-making, enabling organizations to test process changes in a virtual environment without real-world consequences. Despite advancements in automatic simulation model discovery using process mining, BPS is still underused due to challenges in accuracy. *Human-in-the-Loop* (HITL) integrates human expertise into automated systems, where humans guide, validate, or intervene in the automation process to ensure accuracy and context. This paper introduces a framework identifying key stages in BPS studies where HITL can be applied and the factors influencing the degree of human involvement. The framework is based on a literature review and expert interviews, providing valuable insights and implications for researchers and practitioners.

1 INTRODUCTION

Organizations have to regularly improve and redesign their business processes to stay competitive in rapidly changing environments (Măruşter and Beest 2009). As markets evolve and new technologies emerge, businesses are required to remain agile, ensuring that their processes align with current demands. Business Process Simulation (BPS) plays a crucial role by providing insights into potential process improvements, allowing organizations to model, analyze, and implement changes with minimal risk (Wynn et al. 2008). This proactive approach helps companies to maintain efficiency and adapt quickly to external pressures or internal shifts. Organizations consider BPS as highly valuable due to its impact on decision-making and process optimization (Dumas et al. 2018). For example, BPS can analyze the outcomes of increasing the number of resources in a process. Additionally, BPS is a valuable tool to test potential impacts of different simulation scenarios without impacting the real world (Van der Aalst 2010). The accuracy of the results obtained from the simulation depends primarily on the accuracy of the *as-is* simulation model, which must closely match the actual dynamics of the process. The process model is defined as a representation of the business process flow (Weske 2019), and the BPS model is a process model enriched with additional parameters, such as the number of cases or instances that arrive, the processing time for executing activities, the resources performing the tasks, and their availability (Camargo et al. 2020).

Traditionally, BPS models are manually discovered through expert input and data-gathering techniques, such as interviews and on-site observation (Măruşter and Beest 2009). While this approach provides a detailed understanding of business processes, it is time-consuming (Măruşter and Beest 2009) and often overlooks rare or exceptional scenarios (Rozinat et al. 2009). Event logs (Marin-Castro and Tello-Leal 2021) are digital records generated by modern information systems, which systematically capture low-level details about executing various business processes within a company (Khodyrev and Popova 2014). Recently, there has been a significant increase in the availability of event logs in organizations (Pourbafrani et al. 2020; Hulzen et al. 2022), which has directed research toward building BPS models from these logs using data-driven simulation techniques (Camargo et al. 2023; Estrada-Torres et al. 2020; Rozinat et al. 2009; Camargo et al. 2020), offering scalability and efficiency. In this paper, we extend beyond traditional manual methods and focus on data-driven BPS techniques that use event logs as the starting point for discovering BPS models.

Despite the recognized advantages of BPS and advancements in automated methods, their adoption remains limited in practice (Hulzen et al. 2020) due to the challenges posed by data quality and the lack of trust in fully automated models (Hulzen et al. 2020). Human expertise remains crucial in ensuring the validity and accuracy of these models (Uhrmacher et al. 2024), but relying solely on human input compromises the scalability and efficiency of the process. Therefore, a balanced approach is needed—one that integrates human expertise at key stages of the BPS modeling process without undermining the benefits of automation (Khraiwesh and Pufahl 2025).

Human-in-the-Loop (HITL) is a concept that addresses this challenge by integrating human judgment, decision-making, and intervention into automated processes or systems, especially at key stages of BPS, such as objective setting, model validation, and result interpretation. This approach is widely applied in fields such as artificial intelligence (Memarian and Doleck 2024) and machine learning (Gómez-Carmona et al. 2024). In the context of BPS, applying HITL can significantly enhance trust in simulation models while retaining the benefits of automation—especially in managing data quality issues and making critical decisions by targeting interventions at key stages, ensuring a balanced integration of automation and manual oversight. However, achieving this balance between automation and human involvement remains a critical challenge in data-driven processes (Beerepoot et al. 2023). The first step toward achieving an optimal balance is to identify the key stages where human expertise can be integrated into BPS (Khraiwesh and Pufahl 2025). Subsequently, the factors that determine when and where human involvement should be included must be clearly defined. Yet, a comprehensive study that identifies these key stages is still missing. This gap exists due to the limited number of reviews in the BPS field (Rosenthal et al. 2018) and the lack of application of the HITL concept in studies related to BPS. While several studies propose methods requiring some form of human involvement, they rarely use the HITL keyword explicitly (Niloofar et al. 2023). Additionally, these studies tend to define human input at stages specific to their algorithm, domain, and scope. As a result, there is no general framework that provides a comprehensive view of the key stages of the HITL concept in BPS. To address this gap, the study targets the following research questions.

- **RQ1:** What are the stages of a BPS study where human input can enhance the accuracy of the simulation model?
- **RQ2:** What factors determine the degree of human involvement in a BPS study?
- **RQ3:** What are the main obstacles to applying BPS in real-world settings?
- **RQ4:** What are the key research directions and opportunities for advancing BPS model development and addressing current gaps in the field?

In this study, the terms *human input*, *human involvement* and *human participation* are used interchangeably to refer to the HITL concept, which indicates the integration of human expertise alongside automated methods at various stages of BPS. Additionally, it is important to note that normally terms such as *'data-driven simulation'* and *'human-in-the-loop'* are used across different domains with varying meanings. In this paper, these terms are contextualized within the BPS domain, with definitions aligned to its modeling, automation, and decision-support functions. The rest of the paper is structured as follows: Section 2 provides the background and the related work. The research method is described in Section 3. Section 4 presents our findings, followed by a discussion and conclusion in Section 5.

2 BACKGROUND

2.1 Data-Driven Simulation

Simulation helps decision-makers assess model behavior under various conditions (Camargo et al. 2020). *Data-driven simulation* refers to the use of event log data to automatically create or refine discrete-event simulation models. BPS is a data-driven approach that leverages historical event logs to simulate and analyze business processes. Lazarova-Molnar and Li (2019) reviewed existing research on different data-driven simulation methods and explored how data-driven simulation models are developed. Following

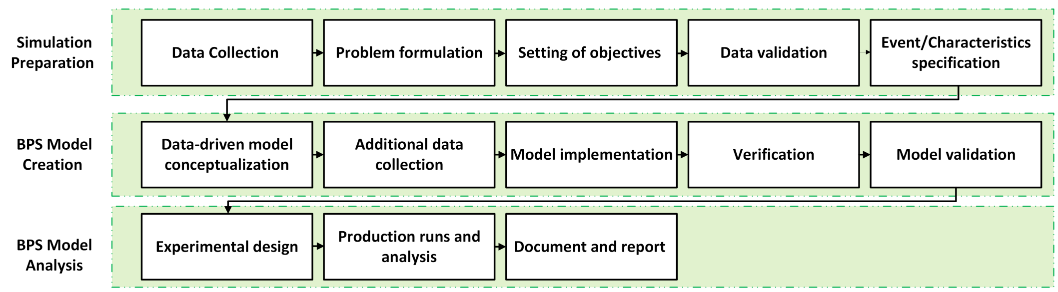

Figure 1: Steps of a data-driven simulation study as outlined by (Lazarova-Molnar and Li 2019) are shown in white blocks. The green areas highlight the phases within the BPS domain.

Lazarova-Molnar and Li (2019), the data-driven simulation study involves 13 steps (white blocks shown in Figure 1). While these steps are directly adopted from Lazarova-Molnar and Li (2019), we introduce a higher-level categorization into three main phases—*simulation preparation, BPS model creation, and BPS model analysis*—to enhance clarity and facilitate explanation within the BPM context (highlighted in green in Figure 1).

The *simulation preparation* phase defines the objectives and the scope of the simulation. It involves collecting event logs, which often contain data quality issues that require cleaning in the data validation step. The next stage is defining the simulation scope to establish a foundation for the simulation, ensuring that all critical elements are considered for a meaningful and actionable model. In the *BPS model creation* phase, a conceptual BPS model is developed, typically using process mining techniques (Van der Aalst 2012) to extract process behavior from event logs. *Process mining* is a technique for discovering process models and extracting insights from event logs. It enables the creation of simulation models by capturing actual process behavior recorded in system data (Van der Aalst 2012). As a key data-driven approach, it supports the construction of as-is BPS models by identifying components like activity duration distributions directly from event logs. Once the conceptual model is enhanced and validated, it is transformed into an executable simulation model, refining its theoretical structure into a form suitable for simulation.

After implementation using tools such as Scylla (Pufahl et al. 2018) and BIMP (Abel 2011), the executable model undergoes verification and validation to ensure accuracy. The *BPS model analysis* phase includes designing experiments, running them to analyze the results, extracting insights, and then documenting the reporting.

2.2 Related Work

The increasing availability of event logs in information systems (Hulzen et al. 2020) has recently driven a growing focus on data-driven BPS. However, as noted by (Rosenthal et al. 2018), the lack of a comprehensive literature review on BPS constitutes gaps in understanding the interrelations and evolution of existing approaches. In BPS area, a literature review by Rosenthal et al. (2018) analyzed papers published between 1999 and 2016, providing valuable insights into application purposes, simulation objectives, challenges in applying BPS, and the role of modeling languages. However, it does not address the HITL concept in the BPS study. Khraiwesh and Pufahl (2025) conducted an SLR on the design of BPS models, analyzing human involvement in BPS development. Still, their study categorizes human involvement into three broad phases: before, during, and after building the BPS model, without details where human input is essential.

The role of human involvement in various data-driven techniques has been explored in multiple studies. For example, Schuster et al. (2022) provide a review on utilizing domain knowledge in process discovery. In contrast, Agrawal et al. (Agrawal et al. 2023) examine the roles Digital Twins (DTs) can play when working alongside humans and assess the extent to which these roles can be automated. Additionally, (Niloofar et al. 2023) highlights the importance of explainability in automated and interactive decision-making processes, proposing an initial framework that identifies key points in the feedback loop of a cognitive digital twin where human involvement can be integrated. However, none of these studies specifically focus on the key stages of HITL in BPS.

This study aims to introduce a framework that identifies where HITL is essential in BPS and defines the factors that determine how extensively it should be applied. We focus on BPS, which constructs *'as-is'* models from historical data stored in information systems.

3 RESEARCH METHOD

To explore where and to what extent HITL can be applied in the development of BPS models and to get further insights into challenges of applying BPS in reality, we conduct a review on existing literature to identify at which stage(s) HITL concept is currently incorporated in BPS methods and tools. We support the review findings with insights from experts by conducting interviews to gain deeper insights into BPS application and the key stages where human involvement is essential within a data-driven simulation study.

3.1 Identifying Key Stages for Human Involvement BPS methods and tools from Existing Research

The *HITL* keyword is not commonly used in BPS. Many studies propose approaches that include human intervention but do not explicitly use the term HITL. To gain an overview of how the HITL concept is applied in BPS, we conducted a dedicated review to identify the key stages with human involvement.

Following the Okoli and Schabram (2015) method, the first phase of the review involves defining the *objective*, which is to analyze the positions and the key stages of human input in a data-driven simulation study. Next are the *protocol and training* steps, where we define the inclusion and exclusion criteria. The search was conducted using three databases: *Scopus, ACM Digital Library, and SpringerLink*. The search focuses on specific terms related to human input, including *"interactive," "incremental," "domain knowledge," "prior knowledge," "hybrid intelligence," "human-in-the-loop"*. Additionally, we included *process mining* and *simulation* to target data-driven simulation, as process mining is the. Process mining is the state-of-the-art data-driven technique used in the BPM field to discover simulation models from event logs (Van der Aalst 2010). The final search term used was:

> *"process mining" AND "simulation" AND ("interactive" OR "incremental" OR "domain knowledge" OR "prior knowledge" OR "hybrid intelligence" OR "human-in-the-loop").*

Table. 1 shows the number of papers obtained from the primary search conducted in February 2025. The search terms were applied to abstracts and, where available, additional metadata. Filters such as language, subject area, content type, and discipline were also used depending on the database.

Table 1: Databases, filters, and number of studies from the February 2025 primary search.

Database	Filters used	#
ACM DL	Search term applied to paper abstract filter	189
Scopus	Search term applied to paper abstract filter	27
Springer Link	Search term applied to paper abstract filter; Language: English; Subject: "Business Process Management"; Content type: Conference paper, article, research article; Discipline: "Computer science"	163
Total		**379**

For the practical screening, we defined a set of inclusion and exclusion criteria to ensure the relevance and quality of the selected studies. We included only studies that introduce data-driven simulation methods (IN1), tools (IN2), or case studies applying a BPS data-driven method or tool (IN3) and incorporate human input in at least one of the data-driven stages (IN4). Only papers published in English (IN5) in a journal or conference proceedings (IN6) are included. Papers classified as tutorials (EX1), surveys or reviews (EX2), fully automated methods or tools (EX3) were excluded. Also, studies not related to BPS (EX4) were excluded.

After applying the defined criteria, an initial screening of 379 identified studies (cf. Table 1) was conducted based on their titles and abstracts, resulting in the selection of 15 studies for further evaluation. In the second round, a full-text review of these 15 primary studies was performed, leading to the exclusion of 7 studies. This left eight studies on which a backward and forward search was carried out, identifying eight additional relevant papers, leading to a total of 16 primary studies for the final review.

In the data extraction phase, we applied a deductive approach, focusing on identifying the type of human involvement described in the data-driven simulation studies. Specifically, we analyzed the stages defined by (Lazarova-Molnar and Li 2019) (outlined in Figure 1) to determine where different forms of human participation were mentioned and assigned them to the relevant steps in the simulation process.

3.2 Supporting Insights from Expert Interviews

In this phase, we conducted online interviews with five experts, three from industry and two from academia. Experts were selected based on their applied experience with BPS projects in academic and industrial settings. Despite the small sample size, the diversity of their backgrounds offered valuable perspectives aligned with our conceptual focus. The interviews were recorded, transcribed, and subsequently coded for analysis. On average, each interview lasted around 45 minutes. The experts represent potential users of data-driven simulation and researchers in the BPS field. In individual, semi-structured interviews, we introduced the data-driven simulation steps outlined in Figure 1. The experts were asked to rate the necessity of human involvement at each stage in Figure 1 on a scale from 1 (not important) to 5 (very important). They were then questioned about the factors affecting the degree of human involvement and the stages they believe are under-researched. Additionally, the experts were asked to identify the obstacles to applying BPS in practice, the challenges of using fully automated BPS, and whether they believe human involvement is essential for improving the accuracy of BPS models. Details are provided in the supplementary material (https://figshare.com/s/79c61613baf239beb472).

4 FINDINGS

In the following, we present our findings from the review based on the data extraction process described above as well as the results of the experts interviews. A full version of the concept matrix and further details for the interviews is available online in the supplementary material.

4.1 Findings from the Review

Based on the results from the literature review, we will present the stages that involve or require human input, starting with those stages that are most frequently mentioned in the studies as needing human input and progressing to those with less frequent human involvement.

We observe that most of the studies (14 out of 16) (Carmen et al. 2015; Ibrahim et al. 2017) include humans in the *data-driven model conceptualization* stage, where additional input from humans is required alongside event logs to build the BPS conceptual model; this input may include simulation configurations (Carmen et al. 2015; Hulzen et al. 2020; Back et al. 2020), assumptions through documentation (Hulzen et al. 2022; Kovalchuk et al. 2018), and missing data gathered through observations (Ibrahim et al. 2017). Additionally, most recent studies (11 out of 16) (Carmen et al. 2015; Ibrahim et al. 2017) include humans in the *Production runs and analysis* stage, where decisions are based on the interpretations of analysts or domain experts. The *model validation* stage also has a strong research focus; more than half of the studies include humans to validate the final model (Pourbafrani and Van der Aalst 2021; Hulzen et al. 2020; Hulzen et al. 2022; Ibrahim et al. 2017; Carmen et al. 2015; Back et al. 2020). Other stages, such as *data validation, problem formulation, setting of objectives, additional data collection, verification and model implementation*, have less research focus, with fewer than half of the papers addressing these stages, with counts of 6 (Kovalchuk et al. 2018; Hulzen et al. 2020; Back et al. 2020), 2 (Ibrahim et al. 2017), 2 (Ibrahim et al. 2017), 3 (Abo-Hamad and Arisha 2013), 1 (Ibrahim et al. 2017) and 3 (Hulzen et al. 2020) out of 16, respectively. Meanwhile, stages such as data collection, documentation and event and characteristics specification are not clearly mentioned with human involvement in any of the studies; this is likely because the primary focus of these studies is on the technical development and execution of BPS models, rather than on documenting the full lifecycle of a BPS study

4.2 Insights from Experts' Interviews

The results obtained from the experts' interviews, detailed in the supplementary material, indicate that the *setting of objectives* is the stage requiring the most human involvement (rated 23 out of 25), followed by *problem formulation*, *model validation*, and *experimental design*, all rated 21 out of 25. According to the ratings provided by the experts during the interviews, these stages are considered essential for human involvement and are less likely to be replaced by automated methods. The *data-driven model conceptualization* and *production runs and analysis* stages follow, with ratings of 19 out of 25, suggesting that while important, these stages could be partially automated while still benefiting from human input. The next group includes the *document and report*, *data validation*, and *additional data collection* stages, with ratings of 18, 16, and 16 out of 25, respectively, indicating a moderate need for human involvement with more room for automation. Finally, the *verification*, *event specification*, and *data collection* stages are ranked with ratings of 14, 14, and 13, respectively. *model implementation* received the lowest score of 12 out of 25. These results show that human involvement is necessary across all stages, though the degree of involvement varies, as reflected in the expert ratings.

4.3 Comparative Analysis

Having analyzed the results from the literature review and expert insights, we applied min-max normalization to enable a meaningful comparison despite differences in scale (Henderi et al. 2021), as shown in Figure 2. The review findings show how often HITL is mentioned in each BPS stage, while expert ratings reflect the perceived importance of human involvement. For example, in Figure 2, the "setting of objectives" phase has a normalized literature score of 0.14, indicating it is rarely addressed in the reviewed papers. In contrast, the normalized expert rating is 0.81, showing that experts consider it highly important. Inductively, based on the comparison of the normalized ratings, we created three categories to reflect the varying levels of human involvement across these stages. Referring to Figure 2, we observe a significant difference between the results, particularly in the stages of *problem formulation, setting objectives, experimental design, and documentation*. We hypothesize that this difference arises from how BPS is represented in research studies. Many papers do not explicitly mention how these steps are conducted; in other words, many studies assume that the problem and objectives are already defined at the outset, especially when focusing on the development or application of a specific method or framework. Similarly, experimental design and documentation are often not emphasized, as they are viewed as standard procedures or background tasks. Researchers typically prioritize novel results and methods, assuming that experimental design follows conventional practices and that documentation is supporting rather than central to the study's contribution. As a result, these stages are often treated as prerequisites or routine steps, with a limited focus on the publication. So, for these stages, we will consider the results obtained from the experts' interviews to define the level of necessity for human involvement. Based on this inductive comparison, *problem formulation, setting objectives, experimental design* stages are classified into **High Human Involvement** category. while *Document and report* is classified to **Moderate Human Involvement**. Furthermore, *Model validation, data-driven conceptualization, production runs, and analysis* stages are also classified into **High Human Involvement** because they are considered essential stages for human involvement, according to experts, and they also receive significant attention in research. Research focusing on these stages indicates that some tasks require human input and are difficult to replace with automated methods. For example, setting simulation parameters often relies on assumptions and contextual factors, which can be challenging to

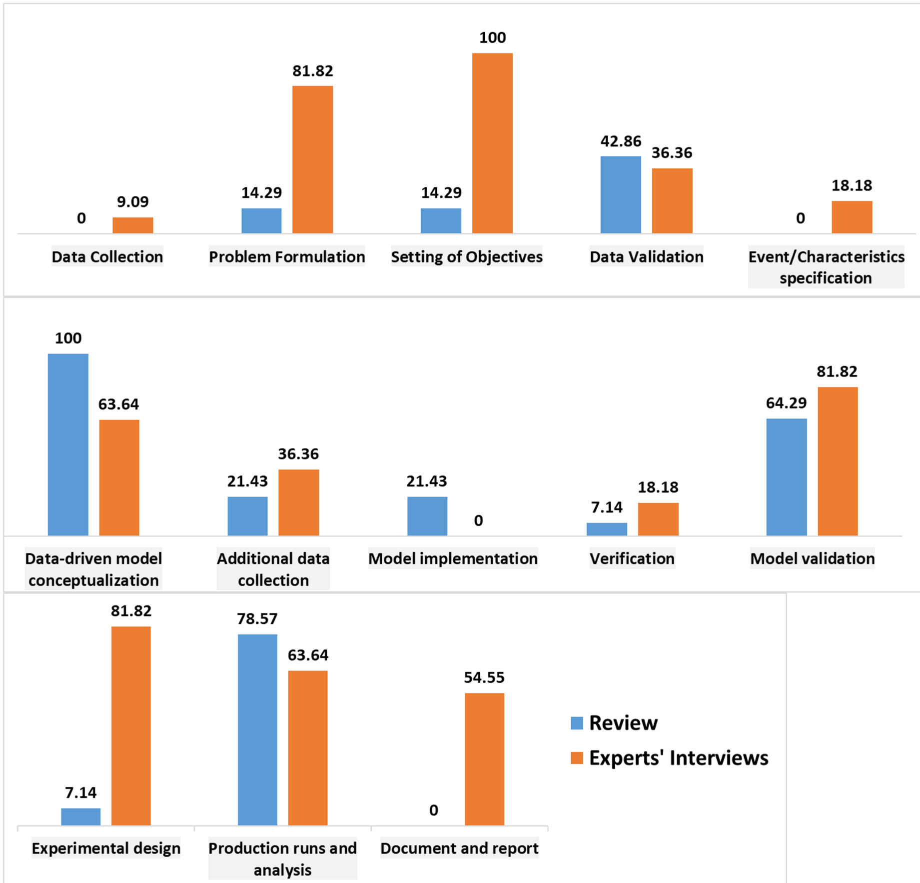

Figure 2: Results of the comparison after applying min-max normalization

capture or obtain through automated methods, as noted in discussions of the complexity of data preparation and the 'semantic gap' in business process simulation (Rosenthal et al. 2018).

In contrast, *data collection, verification, model implementation, and event/characteristics specification* is classified into **Low Human Involvement** since they are the stages that received the least attention in the studies and in the experts' ratings, which indicate that automated methods can do many of the tasks at these stages, require the least human involvement. Meanwhile, *data validation, additional data collection* stages have moderate-rated values in terms of human involvement studies and in the experts' ratings, so they are classified into **Moderate Human Involvement**.

4.4 Resulting Framework

Based on the analysis of our results, we developed a theoretical framework highlighting the key stages where human involvement can take place to answer research question R1. Examining the results from our expert interviews, we observe that the lowest rating was given to the model implementation stage, with a score of 12 out of 25. This indicates that human interaction can be present in all stages, but the degree of necessity varies. Figure 3 illustrates the framework, referring to the steps in Fig. 1, with color codes representing the three categories that indicate the need of human input: black highlight represents **High Human Involvement** with a high necessity for human involvement, white highlight represents **Low Human Involvement** with minimal involvement required, and gray highlight represents **Moderate Human Involvement** reflecting a moderate degree of necessity.

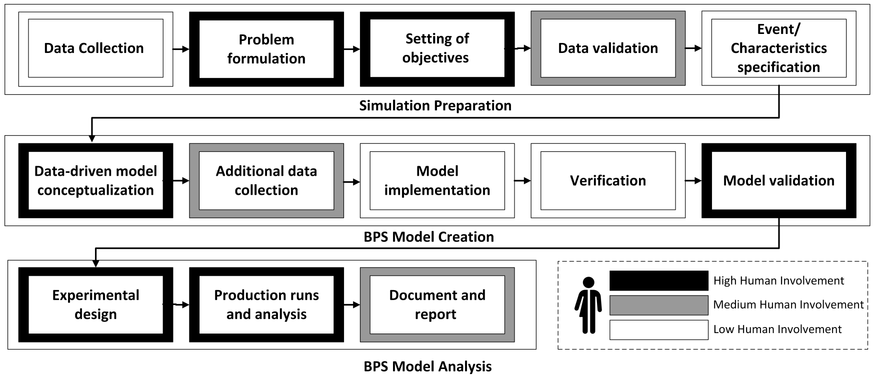

Figure 3: Framework illustrating the key stages of human involvement in BPS with color codes indicating the necessity level: black for high, gray for moderate, and white for minimal involvement.

4.5 Factors affecting the degree of human involvements

Experts provided insights into the factors that affect the degree of human involvement required to answer research question RQ2. The first factor is *data quality*. The starting point of data-driven approaches are event logs, which can significantly impact the accuracy of BPS models. When poor data quality is present, a high degree of human involvement is required to correct or improve the data. An expert mentioned that *"data quality also plays a huge role. Data quality, in the sense of the source of available data, is crucial"*

These quality issues limit the potential of solely relying on process execution data when developing a simulation model in real-life settings (Hulzen et al. 2022). While certain issues, such as negative activity durations or missing data, can be easily detected, others are more challenging to identify without specific domain knowledge, such as determining whether an activity duration falls within a feasible range (Hulzen et al. 2022). The second factor is the *field of study* and the *objective of the simulation*. Simulation studies in sensitive fields like healthcare necessitate a high degree of human involvement to ensure the accuracy and safety of outcomes. For instance, simulation-based approaches have been widely employed to support decision-making in healthcare processes, emphasizing the critical role of human expertise in managing complex systems (Ruiz et al. 2024); in this regard, an expert from academia mentioned *"in healthcare or administrative processes, where many people are involved, the complexity goes up. That means more assumptions need to be made, which increases the need for human involvement."*

Conversely, simulations intended for less critical applications, such as animating processes (Chen 2009; Cramer and Kastens 2009) or illustrating production trends, typically require less human involvement. One example given by an expert from industry; *"If you're looking for general trends rather than specific, detailed cases, human involvement can be reduced. However, human input is necessary for special scenarios not covered by data. The extent of human involvement depends on the objective and the problem definition for the simulation."* Advancements in simulation software have facilitated the creation of detailed 3D models and animations, enhancing visualization and understanding of production processes with minimal human intervention. The third factor is the *complexity* of the simulation and the *skills of employees* who work with it. Dealing with complex simulation situations manually requires highly skilled experts in the field. As the complexity of a simulation increases, so does the necessity for expertise in model interpretation, parameter tuning, and result validation. Studies highlight that digital assistance can enhance performance in complex tasks, but its effectiveness depends on the skill level of the employees (Keller et al. 2021).

4.6 Challenges in Business Process Simulation and Key Insights

During the interview with experts, we investigated the obstacles of applying BPS in real-world scenarios to answer research question RQ3. The experts emphasized that the most significant challenges include the lack of expertise and the inherent complexity of simulation. One expert from academia mentioned *"If you want to conduct a simulation study, you need people who know how to do it. In many organizations, there are no individuals who are truly familiar with conducting simulations. Of course, there are tools that support simulation, but you still need knowledge of simulation and some kind of experience or training."*. Additionally, BPS is often viewed as a complex system, where processes cannot be separated from external influences, expert from industry mentioned *"I think it it's working in that type of isolation consideration. However, business processes are more like a web"*. As a result, practitioners are faced with the dilemma of either making assumptions early on or including all influencing factors, which further compounds the complexity, as clarified by an expert from academia: *"To develop a model that is sufficiently realistic and a valid representation of a real-life process, it also takes a lot of time. Otherwise, if you want to do it quickly, you'll have to make a large number of simplifying assumptions, which would then lead to a model that is no longer realistic, and hence the results are not reliable."*

Another significant concern is trust in simulation results, as event log quality issues (Dakic et al. 2023) limit the reliability of fully automated methods, in addition to challenges related to event log availability. An expert from academia mentioned describing the effect of data quality: *"Data quality is also definitely a challenge in the sense that it could be that your data is not a true reflection of what happens in reality."*

Furthermore, we consulted experts regarding the stages that are under-researched in terms of human involvement. They identified *model implementation, event/characteristics specification, additional data collection, and documentation* as critical gaps; these stages also have less frequent human involvement according to the findings from our review, which further supports these conclusions

5 DISCUSSION AND CONCLUSION

The resulting framework and the insights from the literature review and the expert interviews lead to several important implications for future research and practice in BPS to answer research question RQ4. **First**, data quality issues significantly impact the reliability of BPS studies, resulting in a heavy reliance on manual methods. These methods, while effective, are time-consuming and constrained by the limited availability of domain experts, as discussed by (Măruşter and Beest 2009). Therefore, data validation step requires increased attention in research. New methodologies should be developed to address data quality issues, with a particular focus on integrating human involvement at this stage to ensure high-quality data input, thereby improving the reliability of subsequent stages. **Second**, human involvement remains critical across all stages of data-driven simulation. While some stages are well-explored, others, such as *event/characteristics specification, additional data collection, and documentation*, remain under-researched. There is a need for further investigation into these areas to ensure that human expertise can be integrated effectively to support the simulation studies. **Third**, to overcome the challenges of simulation complexity and the limited number of available experts, it is crucial to invest in employee training. By enhancing the skills of employees and providing step-by-step guidance, organizations can reduce the reliance on a small number of experts, democratizing the knowledge required to implement and operate simulation systems. **Fourth**, effective visualization plays a crucial role in the communication of simulation results and the clarification of complex situations. Therefore, future research should focus on improving the visualization of simulation processes, ensuring that users at all levels can interpret and act on the results with greater clarity and ease. As mentioned by an expert from academia *it it's very important to make things as visual as possible. Showing things visually to humans always works better than presenting them with tables of numbers*. **Finally**, as suggested by one of the experts, advancements in technology should be leveraged to enhance the efficiency of BPS. While data-driven methods offer significant benefits, an optimal balance between automation and human involvement must be maintained.

Threats to validity A key limitation of this study is the small number of expert interviews (five in total), which may not fully capture the diversity of perspectives across the broader BPS community. We acknowledge that this may affect the generalizability of our findings, particularly regarding human involvement in BPS. Additionally, the variability in expert responses—evident through a relatively high standard deviation—reflects differing views among professionals in the field. To mitigate these limitations, we adopted a triangulated approach that integrates both the literature review and expert input. Specifically, for the first research question, we combined insights from existing studies with expert ratings on the relevance of human involvement at each stage. For the second question, which focuses on identifying factors affecting the real-world application of BPS, we extracted the most frequently mentioned barriers across expert interviews and validated them by aligning with findings from existing research. This method allowed us to ground expert opinions in the literature, minimize subjectivity, and strengthen the validity of our results despite the limited sample size.

To conclude, we have reviewed 16 papers out of 379 and conducted five interviews with experts who have either researched the field or have practical industry experience. The results led to the development of a framework that identifies key stages where human involvement is required, categorized by levels. Factors such as *data quality*, *simulation complexity*, *simulation objectives*, and *field of study* determine the necessity of human involvement in BPS studies. Obstacles to applying BPS in reality include simulation complexity, which requires numerous assumptions, data quality issues, and a lack of experts. Future work aims to explore emerging technologies and frameworks that can deepen and expand the HITL concept in BPS. Building on the findings of this study, one promising line of work involves investigating how large language models (LLMs) can assist in translating high-level simulation objectives—typically defined by decision-makers or managers—into precise simulation parameters required for BPS execution. This direction also responds to the challenge of relying on a limited number of experts, which can hinder the practical implementation of BPS. Future studies should validate these findings with a larger and more diverse expert group, potentially through structured surveys or broader case studies, to enhance generalizability and statistical strength. In parallel, future research should focus on developing a framework that integrates human involvement across all stages of the BPS lifecycle, with the degree of involvement dynamically adjusted based on necessity. A further avenue is to identify the specific roles that humans can take at each stage, which would help define the required skill sets and corresponding training needed for different stakeholders. This would support more structured and effective collaboration between automated systems and human experts in simulation practice.

REFERENCES

Abel, M. 2011. “Lightning fast business process simulator”. *Master’s thesis. Institute of Computer Science, University of Tartu*.

Abo-Hamad, W., and A. Arisha. 2013. “Simulation-based framework to improve patient experience in an emergency department”. *European journal of operational research* 224(1):154–166.

Agrawal, A., R. Thiel, P. Jain, V. Singh, and M. Fischer. 2023. “Digital Twin: Where do humans fit in?”. *Automation in Construction* 148:104749.

Back, C. O., A. Manataki, A. Papanastasiou, and E. Harrison. 2020. “Stochastic workflow modeling in a surgical ward: towards simulating and predicting patient flow”. In *International Joint Conference on Biomedical Engineering Systems and Technologies*, 565–591. Springer.

Beerepoot, I., C. D. Ciccio, H. A. Reijers, S. Rinderle-Ma, W. Bandara, A. Burattin, *et al*. 2023. “The biggest business process management problems to solve before we die”. *Computers in Industry* 146:103837.

Camargo, M., D. Báron, M. Dumas, and O. González-Rojas. 2023. “Learning business process simulation models: A Hybrid process mining and deep learning approach”. *Information Systems* 117.

Camargo, M., M. Dumas, and O. González-Rojas. 2020. “Automated discovery of business process simulation models from event logs”. *Decision Support Systems* 134.

Carmen, R., M. Defraeye, and I. Van Nieuwenhuyse. 2015. “A decision support system for capacity planning in emergency departments”. *International Journal of Simulation Modelling* 14(2):299–312.

Chen, J. X. 2009. “Continuous animation and simulation”. *Wiley Interdisciplinary Reviews: Computational Statistics* 1(3):333–337.

Cramer, B., and U. Kastens. 2009. “Animation automatically generated from simulation specifications”. In *2009 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)*, 157–164. IEEE.

Dakic, D., D. Stefanovic, T. Vuckovic, M. Zizakov, and B. Stevanov. 2023. “Event Log Data Quality Issues and Solutions”. *Mathematics* 11(13):2858.

Dumas, M., L. M. Rosa, J. Mendling, and A. H. Reijers. 2018. *Fundamentals of business process management*. Springer.

Estrada-Torres, B., M. Camargo, M. Dumas, and M. Yerokhin. 2020. “Discovering business process simulation models in the presence of multitasking”. In *International Conference on Research Challenges in Information Science*.

Gómez-Carmona, O., D. Casado-Mansilla, D. L. de Ipiña, and J. García-Zubia. 2024. “Human-in-the-loop machine learning: Reconceptualizing the role of the user in interactive approaches”. *Internet of Things* 25.

Henderi, H., T. Wahyuningsih, and E. Rahwanto. 2021. “Comparison of Min-Max normalization and Z-Score Normalization in the K-nearest neighbor (kNN) Algorithm to Test the Accuracy of Types of Breast Cancer”. *International Journal of Informatics and Information Systems* 4(1):13–20.

Hulzen, G. V., N. Martin, and B. Depaire. 2020. “The need for interactive data-driven process simulation in healthcare: A case study”. In *International Conference on Process Mining*, 317–329. Springer.

Hulzen, G. V., N. Martin, B. Depaire, and G. Souverijns. 2022. “Supporting capacity management decisions in healthcare using data-driven process simulation”. *Journal of Biomedical Informatics* 129:104060.

Ibrahim, I. M., C.-Y. Liong, S. A. Bakar, N. Ahmad, and A. F. Najmuddin. 2017. “Minimizing patient waiting time in emergency department of public hospital using simulation optimization approach”. In *AIP conference proceedings*, Volume 1830. AIP Publishing.

Keller, T., M. Behling, C. Stockinger, J. Metternich, and K. Schützer. 2021. “Analysis of the influence of process complexity and employee competence on the effect of digital assistance in industrial assembly”. *Production Engineering* 15:1–8.

Khodyrev, I., and S. Popova. 2014. “Discrete modeling and simulation of business processes using event logs”. *Procedia Computer Science* 29:322–331.

Khraiwesh, S., and L. Pufahl. 2025. “Review of Design of Business Process Simulation Models”. In *International Conference on Research Challenges in Information Science*, 452–469. Springer.

Kovalchuk, S. V., A. A. Funkner, O. G. Metsker, and A. N. Yakovlev. 2018. “Simulation of patient flow in multiple healthcare units using process and data mining techniques for model identification”. *Journal of biomedical informatics* 82:128–142.

Lazarova-Molnar, S., and X. Li. 2019. “Deriving simulation models from data: steps of simulation studies revisited”. In *2019 Winter Simulation Conference (WSC)*, 2771–2782 https://doi.org/10.1109/WSC40007.2019.9004697.

Marin-Castro, H. M., and E. Tello-Leal. 2021. “Event log preprocessing for process mining: a review”. *Applied Sciences* 11(22):10556.

Măruşter, L., and N. R. V. Beest. 2009. “Redesigning business processes: a methodology based on simulation and process mining techniques”. *Knowledge and Information Systems* 21:267–297.

Memarian, B., and T. Doleck. 2024. “Human-in-the-loop in artificial intelligence in education: A review and entity-relationship (ER) analysis”. *Computers in Human Behavior: Artificial Humans*.

Niloofar, P., S. Lazarova-Molnar, F. Omitaomu, H. Xu, and X. Li. 2023. “A General Framework for Human-in-the-Loop Cognitive Digital Twins”. In *2023 Winter Simulation Conference (WSC)*, 3202–3213 https://doi.org/10.1109/WSC60868.2023.10407598.

Okoli, C., and K. Schabram. 2015. “A guide to conducting a systematic literature review of information systems research”.

Pourbafrani, M., and W. M. Van der Aalst. 2021. “Interactive process improvement using simulation of enriched process trees”. In *International Conference on Service-Oriented Computing*, 61–76. Springer.

Pourbafrani, M., S. J. van Zelst, and W. M. Van der Aalst. 2020. "Supporting automatic system dynamics model generation for simulation in the context of process mining". In *Business Information Systems: 23rd International Conference, BIS 2020, Colorado Springs, CO, USA, June 8–10, 2020, Proceedings 23*, 249–263. Springer.

Pufahl, L., T. Y. Wong, and M. Weske. 2018. "Design of an extensible BPMN process simulator". In *Business Process Management Workshops: BPM 2017 International Workshops, Barcelona, Spain, September 10-11, 2017, Revised Papers 15*, 782–795. Springer.

Rosenthal, K., B. Ternes, and S. Strecker. 2018. "Business Process Simulation: A Systematic Literature Review.". In *ECIS*, 199.

Rozinat, A., R. S. Mans, M. Song, and W. M. Van der Aalst. 2009. "Discovering simulation models". *Information systems* 34(3).

Ruiz, M., E. Orta, and J. Sánchez. 2024. "A simulation-based approach for decision-support in healthcare processes". *Simulation Modelling Practice and Theory* 136:102983.

Schuster, D., S. J. van Zelst, and W. M. Van der Aalst. 2022. "Utilizing domain knowledge in data-driven process discovery: A literature review". *Computers in Industry*.

Uhrmacher, A. M., P. Frazier, R. Hähnle, F. Klügl, F. Lorig, B. Ludäscher, *et al*. 2024. "Context, Composition, Automation, and Communication: The C2AC Roadmap for Modeling and Simulation". *ACM Transactions on Modeling and Computer Simulation* 34(4):1–51.

Van der Aalst, W. M. 2010. "Business process simulation revisited". In *Workshop on enterprise and organizational modeling and simulation*. Springer.

Van der Aalst, W. M. 2012. "Process mining: Overview and opportunities". *ACM Transactions on Management Information Systems (TMIS)* 3(2):1–17.

Weske, M. 2019. *Business Process Management - Concepts, Languages, Architectures, Third Edition*. Springer.

Wynn, M. T., M. Dumas, C. J. Fidge, A. H. T. Hofstede, and W. M. Van der Aalst. 2008. "Business process simulation for operational decision support". In *Business Process Management Workshops: BPM 2007 International Workshops, BPI, BPD, CBP, ProHealth, RefMod, semantics4ws, Brisbane, Australia*. Springer.

AUTHOR BIOGRAPHIES

SAMIRA KHRAIWESH is a PhD student in the Department of Computer Science at the Technical University of Munich, Heilbronn Campus, Germany, since 2024. Her research interests include business process simulation, process mining, and machine learning. She obtained her Bachelor's and Master's degrees in Computer Engineering from the German Jordanian University, Jordan. Her email address is samira.khraiwesh@tum.de.

LUISE PUFAHL is a professor in the Department of Computer Science at the Technical University of Munich, Heilbronn Campus, Germany. Her research interests include business process management, process mining, operations research, simulation, machine learning, and data-aware methods for handling regulations and promoting sustainability in business processes. She has applied her research to industries such as healthcare and logistics. she studied Information Systems at the Berlin School of Economics and Law and Humboldt Universität zu Berlin, before completing her doctorate at the Hasso Plattner Institute at the University of Potsdam. She has held postdoctoral positions at Technische Universität Berlin and has been a leader of the software and business engineering research group as an interim professor in 2022. Since 2023, she has been a faculty member at TUM, Campus Heilbronn. Her email address is luise.pufahl@tum.de and her website is https://www.professoren.tum.de/en/pufahl-luise.