Charlie and the Semi-Automated Factory: Data-Driven Operator Behavior and Performance Modeling for Human-Machine Collaborative Systems

Eunji Park
KAIST
Korea
eunji.park@kaist.ac.kr

Yugyeong Jung
KAIST
Korea
yugyeong.jung@kaist.ac.kr

Inyeop Kim
KAIST
Korea
inyeop@kse.kaist.ac.kr

Uichin Lee*
KAIST
Korea
uclee@kaist.ac.kr

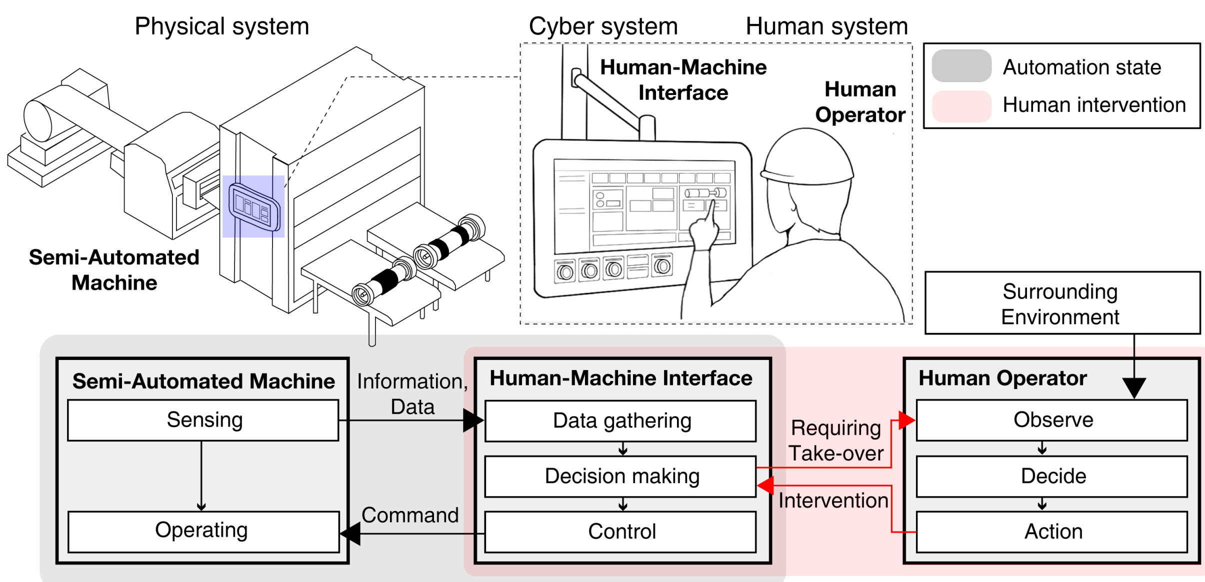

Figure 1: Human-Machine Collaboration Overview in a Semi-automated Manufacturing System. A machine operates in an automated state in a normal situation (gray-box). The human-machine interface (HMI), which facilitates communication between the operator and the machine, allows the operator to know the machine status. When the machine encounters problems, it asks a human to take over a task (red-box). After checking the trouble through the HMI, the operators handle the situation based on their decision

*Corresponding author

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04…$15.00
https://doi.org/10.1145/3544548.3581457

ABSTRACT

A semi-automated manufacturing system that entails human intervention in the middle of the process is a representative collaborative system that requires active interaction between humans and machines. User behavior induced by the operator's decision-making process greatly impacts system operation and performance in such an environment that requires human-machine collaboration. There has been room for utilizing machine-generated data for a fine-grained understanding of the relationship between the behavior and performance of operators in the industrial domain, while multiple streams of data have been collected from manufacturing machines. In this study, we propose a large-scale data-analysis methodology that comprises data contextualization and performance modeling to understand the relationship between operator behavior and performance. For a case study, we collected machine-generated data over 6-months periods from a highly automated machine in a large tire manufacturing facility. We devised a set of metrics consisting of six human-machine interaction factors and four work environment factors as independent variables, and three performance factors as dependent variables. Our modeling results reveal that the performance variations can be explained by the interaction and work environment factors (R^2 = 0.502, 0.356, and 0.500 for the three performance factors, respectively). Finally, we discuss future research directions for the realization of context-aware computing in semi-automated systems by leveraging machine-generated data as a new modality in human-machine collaboration.

CCS CONCEPTS

• **Human-centered computing** → *Collaborative interaction.*

ACM Reference Format:
Eunji Park, Yugyeong Jung, Inyeop Kim, and Uichin Lee. 2023. Charlie and the Semi-Automated Factory: Data-Driven Operator Behavior and Performance Modeling for Human-Machine Collaborative Systems. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3544548.3581457

1 INTRODUCTION

Industry 4.0, the fourth stage of industrialization, has been fueled by digital transformation, and data-driven intelligence in manufacturing has accelerated the so-called "fourth industrial revolution" [34]. Manufacturing processes that were previously performed in an analog manner have been digitized and highly automated, making the processes more efficient. However, many existing automation systems still require intermittent human intervention during the manufacturing processes when the system meets complex errors or sub-tasks that are difficult to automate [29]. For example, production processes such as semiconductor manufacturing or automobile assembly, require human-machine collaboration. Such semi-automated systems that require human intervention are examples of human-machine collaborative systems.

In the human-machine collaborative system, the role of a human operator in a semi-automated system is crucial for ensuring machine automation and work productivity [12, 26, 33]. As illustrated in Figure 1, a machine operates in an automated state in a normal situation (gray-box in Figure 1). The operator can receive information regarding the manufacturing progress via the human-machine interface (HMI) as a cyber component that interfaces communication between the operator and the machine. When the machine encounters a problematic situation that requires human intervention, it asks a human to take over a task (red-box in Figure 1). Subsequently, the operators send the commands back to the HMI which represent the actions based on their decision for handling situation after checking the machine status or trouble through the HMI. The machine then resumes the follow-up tasks after resolving the problematic situation with the aid of the HMI input from the operator. In human-machine collaboration, the decision-making of operators (i.e., how to deal with the current situation by manipulating the HMI) influences the performance (e.g., the production volume and product quality). This is because human behavior in the context of human-machine collaboration may be affected by their cognitive skills or work knowledge, or by environmental factors such as the machine's condition, temperature, and humidity. Therefore, researchers have been interested in interactions between humans and machines in human-machine collaboration systems and how such interactions affect the performance of systems.

The human-computer interaction (HCI) community has been actively exploring data-driven human behavior modeling for optimizing user interactions, such as understanding human routine behaviors (e.g., driving behaviors) [9, 10], mining click-stream patterns on the web [56] and building the data-driven persona of online users [65]. Although understanding human behaviors is also crucial in human-machine collaboration, there is a lack of data-driven, fine-grained analyses of human behavior and work performance modeling using machine-generated data. In the current semi-automated manufacturing machines, multiple streams of machine data (e.g., the manufacturing execution data, detailed steps, material states, and an operator's troubleshooting sequences) are collected via the Internet of Things systems. However, rich machine data have been used in limited contexts to date such as anomaly and security breach detection [5, 30, 52, 57]. We argue that there is room for utilizing machine-generated data to dissect the interaction between an operator and a machine. For example, mining machine-generated data helps to find reference interactions of skilled operators and offer context-aware intervention guidelines for novice operators. As the first step in employing the machine-generated data as a new modality to understand operator behaviors, we analyze the troubleshooting behavior of operators and identify the relationship between how operators interact with the machine over diverse troubles and manufacturing performance. Our study aims to answer the following research questions:

- RQ1. How can the working behavior and performance of operators be modeled using real-world machine-generated data?
- RQ2. Does the work performance vary depending on the work behavior tendency of the operator?

We answer the above research questions through an on-field case study on a tire manufacturing production line. Owing to the safety-critical nature of the tire industry, an extremely small margin of error is required for many quality criteria, such as the length, balance, and weight of each tire part. Rubber is the main raw material that is used in tire production. Because the properties of rubber are highly influenced by the machine setting and the work environment (e.g., temperature, humidity), it is challenging to work uniformly and causes a lot of machine pauses. Thus, the tire building machines require frequent operator interventions for troubleshooting. Some operators may tend to make preemptive parameter adjustments (i.e., proactive interventions) according to their interpretation of the machine data which is generated during the batch (i.e., one cycle) to prevent the stop of a machine. In contrast, other operators may respond to the machine pause that calls for reactive parameter adjustments (i.e., reactive interventions). Since operators' different work behavior tendencies (e.g., proactive vs. reactive trouble handling) can affect manufacturing performance, understanding how operators interact with machines over diverse troubles is critical in semi-automated manufacturing systems.

To introduce the best practice for our methodology, we collected machine and manufacturing data over six months, from February to July 2021. During this period, a total of 8 operators worked for an average of 25.3 shifts and manufactured 57,141 tires and HMI parameters were manipulated 44,405 times. This dataset includes detailed manufacturing processes, alarm histories occurred during production, HMI manipulations, and product information. Using this dataset, we first devised a data analysis methodology to quantify the behavior patterns of operators in diverse contexts, following which domain knowledge from experts was incorporated for further analyses. This led us to propose 13 factors belonging to the following categories: (1) operator-machine interaction factors, (2) environmental factors as independent variables, and (3) performance factors as dependent variables.

Our analysis of machine-generated data with data contextualization and multilevel linear regression revealed that the performance variations could be explained by the interaction and work environment factors (R^2 = 0.502, 0.356, and 0.500 for the three performance factors, respectively). Furthermore, the relevance of each behavior pattern differed according to the type of performance factor. For example, the behavior pattern of responding after trouble occurred was highly correlated to throughput and cycle time. In contrast, the proactive manipulation behavior was strongly related to the alarm recurrence. Furthermore, individual differences in behavior patterns were observed between operators. Considering the findings from the behavior and performance modeling, we call for further research using machine-generated data, such as pattern mining of HMI interaction sequences, context-aware HMIs, a data-driven personal training, and the development of data-driven persona.

2 BACKGROUND AND RELATED WORK

2.1 Human-Machine Collaboration in the Age of Industry 4.0

Human-machine collaboration, which contains interactions between humans and machines, involves the sharing of mutual goals, and the co-management of preemptive tasks [55]. Many recent studies in the manufacturing domain have highlighted the importance of human-machine collaboration because humans and machines can complement one another with their strengths [14, 24, 48, 59]. Human-machine collaboration is necessary for most production processes, such as semiconductor manufacturing lines and automobile assembly processes. Cherubini et al. [14] proposed an approach for collaborative manufacturing with physical human-machine interaction. In such collaboration, humans and automated machines act like teammates [18]. The role of each component in human-machine collaboration may vary depending on the type of task. Endsley et al. [19] classified the automation level of a system into different stages according to the role and degree of human intervention in the machine operation. For example, if humans make decisions and only the machine proceeds with the process, the system is classified as being on the *decision support* level. If the system is involved in both the overall process and decision-making, and humans intervene only when required, it is classified as being on the *supervisory control* level. In recent years, supervisory control has been applied extensively in domains in which the automatic operation is important, such as driving [13, 25], manufacturing [32] and home and medical appliances [50].

2.2 Theoretical Frameworks for Performance Modeling in Industrial Settings

Human intervention, which is dependent on the background knowledge and experiences of the operators, significantly affects the work performance in human-machine collaboration. Moreover, the work performance may be affected by environmental factors or the workplace itself. Various theoretical frameworks have been proposed to relate the work performance (e.g., productivity, errors, quality, and reliability) to various contributing factors such as human factors (e.g., skill and experience) and organizational factors (e.g., rewards and management), under diverse industrial scenarios (e.g., factory assembly lines, nuclear power plant control rooms, and aircraft cockpits).

Researchers have explored the individual status of the operator, which plays an important role in decision-making. Bainbridge claimed that the attributes of the operator's intervention may vary according to their manual control skill, cognitive skills, and long-term knowledge [6]. Researchers have also investigated external factors that could affect the performance of human operators. In an early attempt, Miller and Swain [41] defined performance-shaping factors, i.e., internal factors such as individuals' skill, knowledge, and motivation, and external factors, such as work environment and management, to understand human reliability or the causes and consequences of human errors. Baines et al. [7] synthesized existing approaches in the literature and proposed an integrated human performance modeling theoretical framework in which human performance is defined as a function of individual factors (e.g., personality trait, motivation, skills, and experience), the physical environments (e.g., air quality, noise levels), and the organizational environment (e.g., shift patterns, work teams, job rotation, and communication).

The manner in which humans and machines interact is another important factor since how human operators interact with devices for intervention, or troubleshooting can affect manufacturing performance in a human-machine collaborative system. In particular, in a complex system, human operators may have different work behavior tendencies or troubleshooting strategies for each trouble case according to their knowledge or preference, which may affect overall manufacturing performance. Li et al. [37] emphasized the significance of human-machine interactions that can be captured by human factors (e.g., their capabilities, knowledge, and mental state) and technical factors (e.g., the HMI and operating state). Nachreiner et al. [42] noted that competence in interaction can affect the work performance and operator work-load. These theoretical frameworks for performance modeling comprehensively consider various factors that influence the manufacturing performance. Although human-machine interaction contexts have been established as an important factor, we found that none of the prior studies attempted to tap into continuous streams of machine-generated data to augment the existing framework and analyze the performance-affecting factors at the diverse interaction level (e.g., work behavior tendency level or trouble specific level). This would be of particular interest in semi-automated manufacturing machines, in which tasks are highly automated, and human operators assist in the automation by occasionally handling erroneous machine states or feeding raw materials.

2.3 Data-Driven Analytics in Manufacturing

Manufacturing systems in Industry 4.0 generate a variety of data from interconnected system components. In manufacturing systems, data are continuously generated in the form of signal/information streams (including machine and material state data) of manual interaction data that are entered by operators [35, 43]. The representative standard data collected in smart factories are obtained from a programmable logic controller (PLC), which is frequently used to verify and control parts in networked manufacturing systems. Passively collected PLC data can capture the states of machines and materials, as well as the events or errors that occur during the manufacturing process. Factory workers intervene in manufacturing processes via the HMI in a semi-automated machine, which translates the actions and decisions of workers into PLC commands. As such, the data that are collected on the overall manufacturing situation can be used as a newfound modality for interaction modeling. PLC data enable researchers to uncover new associations, factors, and patterns by aggregating and analyzing the data so that knowledge can be generated for decision support, detecting anomalies, and optimizing prediction models to support key performance indicators (KPIs) [43]. One important challenge for manufacturing systems is the analysis of this massive amount of data, which potentially contain valuable information that is useful for several purposes, such as knowledge generation, KPI optimization, diagnosis, prediction, and feedback for design or decision support [43]. According to Yin et al. [62], the primary purpose of data analysis is to achieve fault-free and cost-efficient execution of a process while realizing the desired level of performance. Nasser and Tariq [45] argued that the critical challenges are (1) integrating and aggregating data types (e.g., machine state and human-machine interaction data) and (2) the extraction and cleaning of information through the extraction of correct data, and the assembly thereof into a suitable structure for data analysis.

Prior studies on HCI have mainly focused on monitoring worker activity in factory settings using wearable sensors. For example, Maekawa et al. [39] proposed an unobtrusive and automatic measurement method for estimating assembly lead time (i.e., the duration of a worker's operation) using a wrist-worn accelerometer to manage and improve the productivity of a line production system. Similarly, Qingxin et al. [47] presented a method for recognizing the assembly work of operators (i.e., starting and ending times of each operation) using data captured by wrist-worn acceleration sensors. To the best of our knowledge, low-level machine state data (e.g., from programmable logic controllers) have mainly been used for computer security [64] and machine fault detection [38] (e.g., detecting malicious code injection or enabling automatic safety vetting) to date. Similar to worker activity monitoring, such machine data analysis merely focuses on recognizing the condition of a machine, without further exploration of human-machine interaction aspects. As stated previously, vast amounts of data can be accumulated from semi-automated machines, and it is of utmost importance to offer real-time data analysis and modeling to assist in human-decision making or interaction guidance in semi-automated manufacturing settings (e.g., what types of user interactions influence the work productivity?) [49, 66].

3 OVERVIEW OF AN AUTOMATED ASSEMBLY MACHINE AND MODELING METHODOLOGY

3.1 Overview of the Tire Building Machine

We used the data collected from a tire building machine, which is considered as a representative semi-automated system. A tire building machine is an elaborate assembly machine that consists of eight automated sub-parts and requires operator-machine interactions, such as material cartridge replacement and error handling. To provide a deeper understanding of the context of the work process, we introduce the basic operations of a tire building machine, including its HMI.

Tire building machines assemble various materials made of rubber and textiles to produce complete tire products. The machines run two tasks in parallel in two separate parts; The first is the *inner drum* that assembles the inner parts of the tire, and the second is the *outer drum* that assembles the outer parts of tire. In addition, the machine consists of hundreds of mechanical parts for winding and assembling the materials, and the material for the next tire is fed right after the previous tire finishes the first step process. Thus, the machine-generated data from the tire building machine includes a complicated history of multiple processes and multiple tires.

The tire building machine contains an HMI that visualizes the process conditions and data to support the real-time monitoring of the machine states. An HMI screen includes touchable elements that enable operators to manipulate the parameters for error prevention and resolution as illustrated in Figure 1. For example, operators can adjust the values of the parameters, such as the pressure or speed that is applied in a particular process. Moreover, the HMI visualizes the trends of the material data, such as the length or weight of the product in progress, so that operators can monitor the data and have a reference for making appropriate troubleshooting decisions.

Tire building machine requires an extremely small margin of error because of safety concerns. So, the machine stops frequently and demands human intervention to maintain performance such as throughput and the quality of products. However, there are multiple ways to resolve troubles due to the complexity of the machine's operation. In other words, there is a lack of standardized methods for troubleshooting. They rely on their background knowledge and work experience to solve problems, so the work performance varies significantly depending on their decisions. Therefore, identifying operators' behaviors (e.g., operator intervention in the machine) is a prerequisite for improving individual operators' performance and further improving the productivity of an assembly line.

3.2 Data-Driven Performance Model Methodology Overview

Figure 2 presents the data analysis methodology for the performance modeling in our study. Each step indicates the five stages of the methodology. For the sake of illustration, we used an enterprise

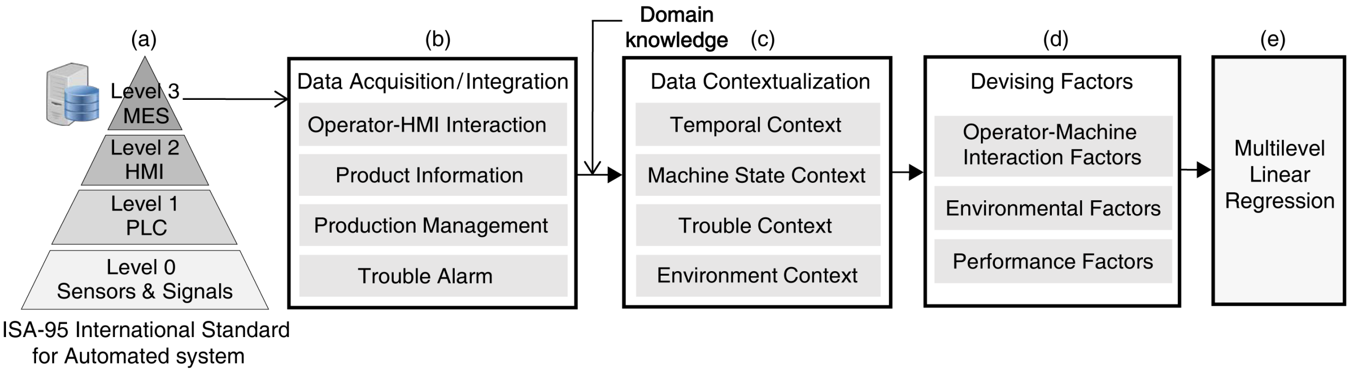

Figure 2: Data analysis methodology. (a) The data acquisition system of the tire building machine follows the ISA-95 standard that comprises sensors and signals, programmable logic controller (PLC), human-machine interface (HMI), and manufacturing execution system (MES) (b) We extracted four kinds of machine-generated data and integrated it. Then we filtered the missing data period (c) We defined contexts that help us understand the situation of operator-HMI interaction (d) We devised 14 factors that can quantify operator-machine interaction, environment, and working performance (e) We used multilevel linear regression model to determine the effect of operator-machine interaction and environment on performance

hierarchy from ISA-95 [4], which is an international standard for developing automated systems. As in Figure 2.a, the sensors in the building machine at level 0 sends signals to the programmable logic controller (PLC) interface at level 1. The data acquisition system at level 2 collects the data from the PLC operations and enables data visualization via the HMI. Thereafter, the collected data are forwarded to the designated manufacturing execution system (MES) servers at level 3.

Then, we extracted four types of data from the MES server (Figure 2.b): (1) the operator-HMI interaction data, which are collected whenever operators manipulate the parameters for error prevention or troubleshooting; (2) the product information data, which are collected when a tire is completed; (3) the production management system (PMS) data, which are collected when each step of tire production is performed; and (4) the trouble alarm data, which are collected every time the machine encounters trouble.

Subsequently, based on the extracted data, the four data types are integrated based on the timestamp. After preprocessing, we defined contexts that helped us to understand the situation of production based on relevant domain knowledge in the data contextualization stage (Figure 2.c). For gathering and organizing the domain knowledge, we cooperated with four experts who have worked for the tire manufacturing company for 15 years in average. As a result, we extracted the temporal, machine state, trouble, and environmental contexts from the data which will be explained in section 4.2. Using the contextualized data, we devised 13 factors to quantify the operator-HMI interaction (six factors), environment (four factors), and working performance (three factors) as indicated in Figure 2.d. We performed multilevel linear regression to determine the effect of operator-machine interaction and environment on performance in the modeling stage (Figure 2.e).

4 DATA ACQUISITION AND CONTEXTUALIZATION

We present the data acquisition system for the tire building machine and the types of data that were collected. Furthermore, we introduce the five types of contexts that were extracted from the dataset.

4.1 Data Acquisition and Integration

Figure 2.a and Figure 2.b depict the process by which the data acquisition system acquires the machine state and human-machine interaction data. Four types of data are stored and extracted from the servers: (1) operator-HMI interaction data, (2) product information data, (3) process step information data, and (4) trouble alarm data. The tire company had been collecting and using data generated by the machine for evaluating and improving the productivity of tire manufacturing, as well as optimizing their semi-automated manufacturing machines. As part of the funded research project from the tire company, we received a pseudonymized dataset for privacy and security reasons, and Section 3 Special Cases Concerning Pseudonymous Data of Personal Information Protection Act in Korea obviates the need of receiving informed consent regarding sharing a pseudonymized dataset. The company's manufacturing department explained to the operators the purpose of data collection and research goals. Since the dataset did not include personally identifiable information, this research was exempted from the IRB review.

4.1.1 Operator-HMI Interaction Data. When a problem occurs in a machine, the machine sends a trouble alarm to the operator through the HMI. This process resembles the situation where an automated driving system asks the driver to take over when the system encounters a tricky situation. Then, the operator check the trouble alarm message through the HMI. The operators can subsequently change the production settings by manipulating specific manufacturing parameters within the allowable ranges. The HMI data that we collected includes the timestamp, parameter names, upper and lower bounds of each parameter, and modification history of the manufacturing parameter values (e.g., values before and after the change).

4.1.2 Product Information Data. Various products can be produced in a factory using the same machine. The types of problems and methods for dealing with the problems may vary depending on the product types being manufactured. Therefore, product information is required to contextualize the situation or background in which the operator intervenes. For example, in this study, the tire-building machines produced tires of various sizes and thicknesses. The product information data contains information on each tire, such as the unique bar code, tire specifications, standard weight, measured weight, working group, and tire production start and end times.

4.1.3 Production Management System (PMS) Data. The PMS is a very basic data in process management that includes the temporal context. Typically, each production step follows a predetermined process sequence consecutively. The tire assembly process consists of 27 steps that involve the assembly of the inner and outer drums. The PMS data records the timestamp for each step along with the process name. Every time a particular step starts and ends, the name of the step along with the start and end times are recorded. Furthermore, the operator name, working group, and specification of the tire are attached to the step information. Pseudonymized data from the company were used for analysis in order to protect the privacy of the operators and the security of the company.

4.1.4 Trouble Alarm Data. When a critical problem occurs, the machine stops and requests the operator to take over the task. In certain cases, the machine notifies the operator that intervention is required by sending a trouble alarm message. Then, the operator solves the problem based on the trouble alarm message by manipulating the parameter of the HMI. As such, a trouble alarm history aids in classifying the circumstances in which intervention is required. The trouble alarm data that we collected includes the timestamp, types, alarm duration, and priority of the alarms.

4.1.5 Data Integration. We integrated the four data (operator-HMI interaction, product information, PMS, and trouble alarm data) based on the timestamps. We unified all time units into millisecond units. Unified timestamps enable the integrated analysis of relevant events by simultaneously referring to different data sources. For example, it is possible to determine which operator worked during which work shift, which production steps were affected, and which HMI parameters were manipulated for a given problem or alarm.

Because the data in this study were collected in the wild, there were missing data collection periods; there were 30 days of missing for trouble alarm data, 54 days of missing for product information data, 47 days of missing for operator-HMI interaction data, and 65 days of missing for product management system data. When some or all four types of data were not collected during the working shift of one operator, possibly owing to a long-term breakdown of a manufacturing machine, or database problems or server outages. Unfortunately, if one data contains a missing value at a certain period, we cannot fully understand the circumstances at that period, even if other data were collected. For this reason, we discarded the periods with missing data. As a result, 68 out of 180 days were used for further data analysis. Additionally, among 12 operators, we exclude 4 operators who temporarily worked only once or twice over the six months. As shown later, the remaining data are still sufficient to identify an operator's behavior and working context. During the remaining periods, operators manipulated 252 parameters 44,405 times, 596 trouble alarms occurred 1,274,034 times, and 57,141 tires were produced.

4.2 Data Contextualization

Subsequently, the raw data were processed for contextualization (i.e., labeling or categorizing raw data for semantic filtering) by incorporating relevant domain knowledge. Contextualization can be used to model the user behavior in an interaction situation [22] by linking the output from users to their surrounding circumstances [58]. The context describes the features of the environment within which a human performs a task [17]. Moreover, the contextualization of data is essential in terms of semantic analysis to provide valuable inputs for analyzing manufacturing data [21].

We built *metadata* for data contextualization by collaborating with four domain experts from the tire manufacturing company. The raw data obtained from the MES server contains a large amount of data, but the raw data itself does not provide useful information. To extract valid knowledge from the raw data, it is necessary to know which event (e.g., HMI manipulations and alarm events) occurred in which contexts (i.e., machine parts, process steps). Therefore, we created metadata that enables *semantically stitching* various data beyond simply stitching based on timestamps. We considered the machine parts (as a spatial connection) and the manufacturing process (as a temporal connection) for semantic stitching. We labeled each event, such as HMI manipulations and alarm events, with the relevant machine parts and the manufacturing processes. In addition, we organized the key elements of the production, including shift, cycle, and machine state. The metadata helped us to contextualize the operator-machine interaction data in terms of the temporal domain, machine state, trouble events, and environmental contexts.

4.2.1 Classifying of Temporal Context. Shift-working systems are common in the manufacturing industry; the operators repeat multiple production cycles during each shift. More specifically, various events such as alarms or HMI parameter manipulations occur during one assembly cycle. The events are combined to form a cycle and the cycles are combined to form a shift. As such, the temporal context can be divided into several levels according to the length of time, thus the extracted contexts may differ depending on the temporal context. In this study, we classified the temporal contexts into three levels: shift, cycle, and event.

The data that we collected for this study included work history from February to July 2021. During this period, the operators worked for 8 h per shift in a three-shift working system. We defined the *cycle* as the time interval for assembling a tire. We defined the initial step in which the materials are fed into the tire building machine as the start of the cycle, and the final step in which the tire is discharged from the machine as the end of the cycle. According to our data analysis, the average time for one cycle was 79.15 s. Finally, as shown in Figure 3, the smallest unit of the temporal context is the event-level context. We considered the specific time of HMI manipulation or alarm occurrence as discrete events.

4.2.2 Defining Machine State and Condition. Automated machines assemble products automatically without requiring human intervention unless problems arise. However, the machine will be stopped if a problem occurs, possibly owing to machine faults (e.g., sensing or mechanical errors) or material issues (e.g., unqualified sizes). The machine can resume automation after the operator implements

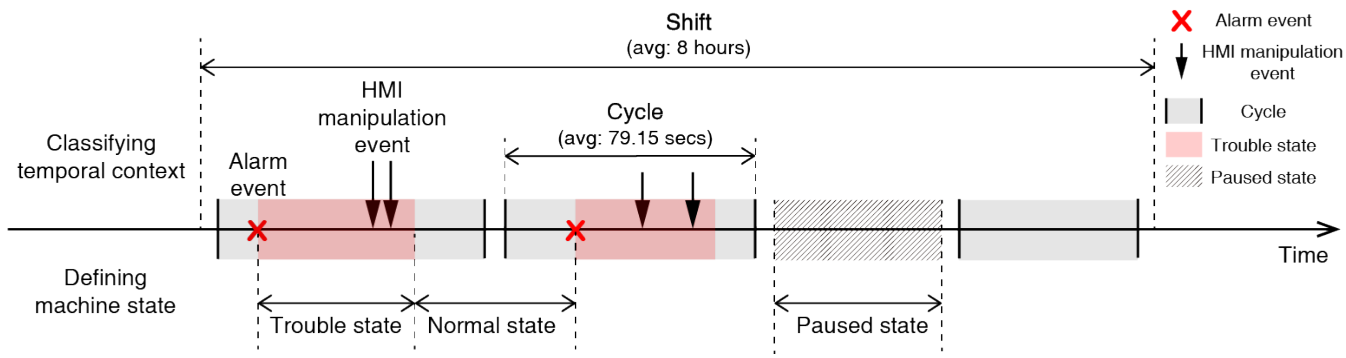

Figure 3: Context extraction: Classifying temporal context and defining machine state. The temporal context is divided into shift (the period an operator works once), cycle (the period for assembling a tire), and event (such as an HMI manipulation or trouble alarm). The machine state is classified into the normal state (the machine is running automatically), trouble state (the machine is stopped owing to problems), and paused state (the machine is paused for reasons such as material cartridge replacement)

an appropriate action for the problem that corresponds to human intervention. Based on the collected data, we divided the state of the tire building machine into three categories: 1) the normal state (in which the process is running automatically); 2) the trouble state (in which the machine is stopped owing to problems); and 3) the paused state (in which the machine is temporarily paused for reasons such as material cartridge replacement, which are unrelated to the problems) as in Figure 3. Therefore, we used the alarm data to classify the machine state. The trouble state time was defined as the time from the alarm start until the alarm was removed from the HMI screen. The paused state was defined according to the PMS data that contains equipment repair and pause history.

4.2.3 Matching of Alarm Events to HMI Parameter Manipulation Events. The operator performs troubleshooting when an alarm occurs. For example, as illustrated in Figure 4, the operator manipulates the HMI parameters to solve the problem. However, it is difficult to infer which alarm the operator is targeting. This is particularly challenging when alarms and parameter manipulations occur continuously over a short period. So, we matched the alarm records and HMI manipulation records based on the metadata.

For matching, we labeled the related machine parts (e.g., inner drum or outer drum) to the HMI parameter manipulation log based on the metadata. Furthermore, we assumed that the operator will act within the same cycle if an alarm occurs. Therefore, we matched the alarm event to the HMI parameter manipulation event in the case of that: 1) the alarm event and the HMI manipulation event take place within the same cycle; 2) the HMI manipulation event happens after the alarm event; 3) and the relevant machine parts of each event are identical.

4.2.4 Categorizing Properties of Each Event. Operators intervene in automated manufacturing processes for various purposes. For example, operators may intervene to solve problems that occurred, or prevent them prior to occurrence. We classified the purpose of each HMI manipulation event based on the matching result to understand the purpose of the operator intervention. As in the Figure 4, we classified each HMI manipulation action as either reactive manipulation (e.g., operators manipulate HMI parameters after trouble alarms occur) or proactive manipulation (e.g., operators manipulate HMI parameters in advance to prevent the occurrence of problems). Moreover, actions that did not match any alarm with unknown purposes were classified as unidentified actions.

4.2.5 Categorizing Working Environment. We also categorized the working environment context not related to the machine state (e.g., seasons and shifts). We identified the seasons and working hours from the data timestamps. We classified the six-month data into two seasons, namely spring and summer. Given that the duration of a shift was 8 hours, the working hours were divided into three shifts: morning (6 AM to 2 PM), afternoon (2 PM to 10 PM), and night (10 PM to 6 PM).

5 FACTOR EXTRACTION FOR MODEL BUILDING

We considered the operator-machine interaction factors based on Li et al.'s work which emphasizes the importance of operator-machine interaction contexts for worker performance modeling [37]. We also considered environmental factors, which may be associated with the conditions of operators, manufacturing line states, or materials. We devised these factors by combining the contexts (e.g., the temporal context and machine state) that were introduced in Section 4.2. Table 1 presents the contexts that were used to produce each factor. Finally, we considered three performance factors to measure different performance aspects.

5.1 Operator-Machine Interaction Factor

HMI manipulation is a representative behavior in operator-machine interaction, and existing differences between operators in handling troubles may affect their performance. In a semi-automated manufacturing system, when a specific problem occurs, an operator makes decisions and troubleshoots based on the information in the HMI. In a complex system, such as a tire-building machine, operators may have multiple solutions for troubleshooting, and

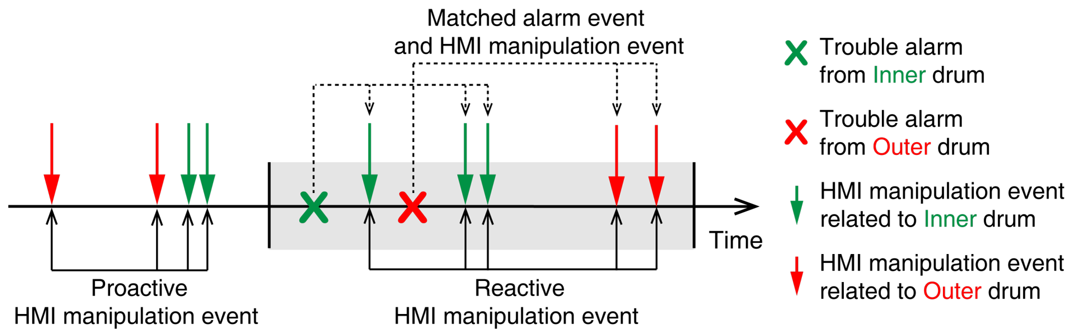

Figure 4: Context extraction: Matching alarm events to HMI parameter manipulation events. The gray box with bold lines indicates one cycle. If an HMI manipulation event occurs within the same cycle after an alarm event, and both are related to the same machine parts, we matched the alarm event to the HMI manipulation event. We classified each HMI manipulation as reactive manipulation (manipulating an HMI after trouble alarms) or proactive manipulation (manipulating an HMI in advance to prevent problems)

Table 1: Combined contexts for factor extraction. We extracted interaction, environmental, and performance factors by combining the contextualized information in Section 4.2 such as temporal context, machine state, matching events, event properties, and working environment.

	Temporal Context	Machine State	Matching Events	Event Properties	Working Environment
Operator-Machine Interaction Factors					
Reactiveness to Troubleshooting	•	•	•	•	
Proactiveness to Troubleshooting	•	•	•	•	
Utilization of HMI for Troubleshooting	•	•			
Responsiveness to Troubleshooting	•	•	•		
Amount of Parameter Modifications	•				
Average Number of Modified Parameters	•				
Environmental Factors					
Number of Spec Changes	•	•			
Number of Alarm Types	•				
Working Time	•				•
Working Season	•				•
Performance Factors					
Trouble Alarm Reoccurring Time	•	•			
Cycle Time	•				
Throughput	•				

they may take actions based on their experience and preference, which could affect the manufacturing performance. In general, the operator-machine interaction can be modeled at diverse levels of detail, ranging from (1) high-level work behavior tendencies to (2) low-level trouble-specific behaviors. Here, work behavior tendency factors describe an operator's general tendency about how the operator interacts with machines over diverse troubles (e.g., proactive vs. reactive trouble handling). In contrast, trouble-specific factors indicate how the operator manipulates the HMI parameters for each specific trouble. Since there is a trade-off between generalizability and specialization, we mainly focused on generalizable factors, which are applicable to other semi-automated manufacturing domains. In addition, modeling trouble-specific factors requires sufficient troubleshooting cases for each alarm, but in reality, there are too many trouble instances in a complex system. In our case, we identified 596 trouble alarm types and 252 unique modified parameters. The distribution of troubles and modified HMI parameters are highly skewed and have long tails. Since there are limited samples, it is challenging to consider trouble-specific factors, which may cause an overfitting problem for the performance model. For this reason, this work mainly focuses on work behavior tendency as operator-machine interaction factors. In our work, each factor is calculated by averaging the results of one shift (i.e., 8 hours) for one operator.

5.1.1 Reactiveness to Troubleshooting. In Section 4.2.4 and Figure 4, we defined an event as a reactive manipulation if operators manipulate HMI parameters after trouble alarms occur. *Reactiveness to troubleshooting* indicates the reactive tendency of an operator in HMI manipulation for trouble handling (i.e., troubleshooting). We calculated the *Reactiveness to troubleshooting* by dividing the number of reactive HMI manipulations by the total number of HMI manipulations for each working shift.

5.1.2 Proactiveness to Troubleshooting. In contrast to *Reactiveness to troubleshooting*, an operator can manipulate HMI parameters in advance to prevent the occurrence of problems. In this case, if the operator appropriately manipulates the HMI in advance, the automation will be better maintained. *Proactiveness to troubleshooting* indicates the proactive tendency of an operator in HMI manipulation to maintain the automation of the machine (i.e., troubleshooting before trouble). We calculated the *Proactiveness to troubleshooting* by dividing the number of proactive HMI manipulations by the total number of HMI manipulations for each working shift.

5.1.3 Utilization of HMI for Troubleshooting. There is another option when a operator encounters the trouble alarm. It is directly manipulating the machine and material manually. For example, to solve problems related to material alignment, an operator can manually handle the material instead of manipulating the HMI parameters that control the pressure level of the machine toward the material. We calculated the *HMI-mediated troubleshooting* by dividing the number of HMI manipulations by the number of trouble alarm occurrences.

5.1.4 Responsiveness to Troubleshooting. After the operators check which trouble alarms have occurred, they decide how to handle these problems and take appropriate action. The duration from trouble occurrence to troubleshooting (i.e., response time) may vary depending on the work experience or characteristics of the operator. A prompt response to trouble is closely related to the performance, because a typical assembly cycle requires a short response time (e.g., approximately 1 min). Therefore, we calculated the *responsiveness to troubleshooting* indicating how quickly operators respond to trouble, by calculating the time interval from the trouble alarm occurrence to the adjacent HMI manipulation matching the trouble alarm.

5.1.5 Amount of Parameter Modifications. Operators should modify HMI parameters to proper values to handle troubles because, otherwise, trouble alarms occur again, stopping the machine. The tire building machine has many inter-connected components, and the tire assembly process is highly complicated. In addition, rubber, the primary material of tires, can shrink or stretch depending on the environment. Thus, it is challenging for experienced operators to modify parameters to appropriate values at once. Amount of parameter modifications may be related to the recurrence of troubles and consequently affect the work performance. We quantify *amount of parameter modifications* by dividing amount of parameter modification by the range of upper and lower threshold for each parameter: $\sum_{i=1}^{n} \frac{m_i}{u_i - l_i}$ (m_i = amount of parameter modification for ith manipulation, u_i, l_i = upper and lower threshold of the manipulated parameter in ith manipulation). After calculating the sum, we averaged the values during one working shift.

5.1.6 Average Number of Modified Parameters per Action. Depending on the circumstance, operators should consider a set of HMI parameters rather than a single parameter for trouble handling due to the complexity of the tire-building process. Based on work experience, some operators may comprehensively modify parameters to solve troubles, while others may consider small sets of parameters or only one parameter. We investigated this aspect by devising the *average number of modified parameters per action* factor that indicates how many parameters operators consider simultaneously for handling troubles. In order to quantify this, we defined a set of matched HMI manipulations (Figure 4) based on the alarm-HMI matching method mentioned in Section 4.2.3 and counted the number of HMI manipulations.

5.2 Environment Factors

Existing studies on work performance have addressed that the work environment affects the operator's performance [7, 11, 12]. By considering available data from our machine database, we considered the following environmental factors: (1) occurrence of resetting manufacturing line, (2) diversity of trouble types, (3) working time, and (4) working season.

5.2.1 Number of Spec Changes. A machine may stop without trouble due to resetting issue in the manufacturing line. In our study, changing the tire specification to be produced is the representative case of resetting the manufacturing line. Stopping a machine to change tire specifications is a typical event in the manufacturing plan. Since stopping a machine does not mean stopping the tire assembly process, it does not affect the tire production cycle time. However, performance factors such as throughput can be affected because the machine can not produce tires while resetting the manufacturing line. We investigate the influence of changing the specifications on performance by counting the number of changes during working hours.

5.2.2 Number of Alarm Types. If the types of trouble alarms are varied, the operator must come up with numerous solutions to address the alarms. Therefore, troubleshooting may be more challenging when various alarms occur sporadically than when the same alarm occurs repeatedly. Based on the expert's opinion, the types of trouble alarms that mainly occur may vary depending on environmental conditions such as temperature and humidity because tires are made of rubber with uneven physical properties. We considered *diversity of trouble types* as an environmental factor.

5.2.3 Working Time. Four teams of operators work in three shifts (i.e., morning, afternoon, and evening) for 8 h. Depending on the operator, there may be preferred working hours. Additionally, depending on the working hours, the number of employees and the effectiveness of communication may differ. We check whether work shift time impacts the operator's performance.

5.2.4 Working Season. Rubber is the primary raw material for tires and is greatly affected by season (e.g., humidity and temperature). Since the state of materials causes troubles in a machine, we check whether the season affects the operator's performance. We note that most of the data in this study was collected during spring and summer.

5.3 Performance Factors

We aim to understand how operator-machine interaction and environmental factors affect the operator's performance. We considered three performance factors mainly used to quantify the operator's performance in the previous studies [54]. Three levels of temporal context are considered for quantifying performance factors (i.e., event-level, cycle-level, and shift-level) as follows: (1) Interval between trouble recurrence for event-level, (2) production cycle time for cycle-level, and (3) production throughput for shift-level.

5.3.1 Trouble Alarm Reoccurring Time. It is essential for operators to properly manipulate the HMI parameters so that the same trouble alarm does not occur again. We calculated the average time interval between trouble recurrence, which indicates how long the same trouble did not occur after an operator solved the trouble. A longer time means that the same trouble did not occur for a long time. We assumed it is related to the fact that an operator appropriately modified the HMI parameters, and the automation of the machine was maintained well. We calculated this factor as follows: $\frac{1}{a} \times \sum_{i=1}^{a} t_i$ where a=the number of alarm events that the same alarm has recurred during the work shift, t_i=the time interval until the same type of alarm as the ith alarm occurs again.

5.3.2 Cycle Time. The time it takes to assemble a tire is the cycle time. If a machine stops due to trouble, the length of the cycle time is affected. As a performance metric, we defined the production cycle time by averaging cycle time during one work shift (8 h). The metric is devised from average flow time [3] which is mainly used in manufacturing studies, which means the average time a unit product is produced.

5.3.3 Throughput. When the machine is stopped due to trouble, the number of tires produced during a given time (i.e., throughput) is reduced. We considered the throughput as a performance factor of how many tires an operator produced during a work shift (8 h).

6 DATA ANALYSIS

First, we determined whether there was a variance in the performance, such as the throughput, production cycle time, and time interval between trouble recurrences. The operators produced an average of 312.33 tires over 8 h (SD = 53.83), and the average cycle time taken for producing one tire was 79.15 s (SD = 12.47). Furthermore, the average trouble alarm recurrence time was 611.29 s (SD = 128.31). We confirmed that the throughput was not constant despite the use of an automated machine.

6.1 Multilevel Regression Analysis

We collected data containing the work histories during a shift of 8 operators. Thus, the data can be clustered into individual operators, and there may be individual differences between the operators. Multilevel models are mainly used when data can be aggregated into higher-level groups (i.e., operators), and there is heterogeneity between groups. In general, the multilevel models allow us to analyze the relationship between dependent and independent variables without the effects caused by the differences between groups, also known as random effects. In this study, we set the performance factors as dependent variables. Interaction and environmental factors (i.e., fixed effects) and operators (i.e., random effects) were set as independent variables. For the mixed-effects models, marginal R^2 indicates variance explained by fixed factors, and conditional R^2 indicates variance explained by both fixed and random factors [44].

6.2 Production Performance Model

The results demonstrate that interaction factors have considerable effects on performance factors, and the performance can be predicted more effectively when considering both the interaction and environmental factors. Interestingly, in light of the fact that the conditional R^2 was higher than the marginal R^2 for all dependent variables, there were significant individual differences in the work-related behavior among the operators (Table 2).

6.2.1 Trouble Alarm Reoccurring Time. The time interval between trouble recurrence measures the interval until the same alarm occurs. We expected that the recurrence time would be longer if the problem was properly addressed appropriately by an operator. The marginal R^2 was 0.376, and the conditional R^2 was 0.502, which means that a higher R^2 value was obtained when considering the individual differences among the operators. Moreover, we confirmed that proactiveness is a significant interaction factor for the alarm recurrence time. As the proactiveness increased by 1%, the alarm recurrence time increased by 1.5 s. No significant factors were related to the HMI manipulation. However, environmental factors such as the number of alarm types, season, and working time affected the alarm recurrence time. The alarm recurrence time was longer when more alarm types occurred. Moreover, the alarm recurrence time was longer during the day than at night, and shorter in summer than in spring.

6.2.2 Cycle Time. The average production cycle time is the average time that is required to assemble a tire. As efficient production is important in manufacturing, the completion of the same process within a short cycle is considered as good performance. Therefore, we assumed that a shorter average production cycle time indicated better performance. The marginal R^2 was 0.253, the conditional R^2 was 0.356, and three factors exhibited a significant correlation with average cycle time. Among the interaction factors, reactiveness to troubleshooting was significant. When the reactiveness to troubleshooting was high, the cycle time was negatively affected. The number of alarm types and working season were significant among the environmental factors. The cycle time was longer when various alarm types occurred and shorter in spring than in summer.

6.2.3 Throughput. The throughput refers to the number of tires that is produced in one shift, which is a measure of production efficiency. According to the throughput results, the marginal R^2 was 0.314 and the conditional R^2 was 0.500. Among the interaction factors, the reactiveness to troubleshooting and amount of parameter manipulation were significant. When both factors increased, the throughput was negatively effected. The number of alarm types and specification change factors were significant among the environmental factors. The throughput was poor if the alarm types

Table 2: Results of multilevel regression on the operator-related behavior. Operator-machine interaction factors and environmental factors are set as independent variables (fixed effects), and operators are set as dependent variables (random effects) (*p<0.05, **p<0.01, ***p<0.001)

	Trouble Alarm Reoccurring Time			Cycle Time			Throughput		
	β	p-value		β	p-value		β	p-value	
Modeling with Interaction Factors and Environmental Factors									
Operator-Machine Interaction Factor									
Reactiveness	-55.43	0.507		25.26	0.008	**	-82.97	0.035	*
Proactiveness	150.12	0.026	*	12.07	0.115		-51.97	0.099	.
Utilization of HMI	8.13	0.812		3.99	0.293		-7.30	0.654	
Responsiveness	-1.01	0.531		0.20	0.278		-0.78	0.305	
Amount of parameter modifications	156.29	0.434		30.28	0.179		-283.28	0.003	**
Avg num of modified parameters	-8.60	0.427		-2.20	0.075	.	6.13	0.227	
Environmental Factor									
Machine-related									
Number of spec changes	13.23	0.181		1.30	0.245		-10.10	0.030	*
Number of alarm types	7.88	0.000	***	0.45	0.000	***	-2.87	0.000	***
Working environment									
Job rotations (afternoon)	41.40	0.014	*	3.70	0.052		-10.16	0.192	
Job rotations (morning)	29.90	0.074		0.83	0.660		-3.57	0.648	
Job rotations (evening)	-	-		-	-		-	-	
Working seasons (spring)	-78.34	0.000	***	-8.25	0.000	***	26.02	0.000	***
Working seasons (summer)	-	-		-	-		-	-	
Marginal R^2	**0.376**			**0.253**			**0.314**		
Conditional R^2	**0.502**			**0.356**			**0.500**		

varied and the specification change was frequent. Furthermore, the throughput was better in spring than in summer.

6.2.4 Relationship between Performance and Interaction Factors. In addition to the multilevel regression analysis, we further investigated the interaction factors that significantly affect the work performance factors (e.g., proactiveness affects trouble reoccurring time, reactiveness affects cycle time, amount of parameter modifications, and reactiveness affect throughput). Through this, we attempted to determine whether the performance of the operator affects the relationship between the independent variable and the dependent variable. Of the eight participants, we selected the top four operators as the high-performance group, and another four operators as the low-performance group.

In the two of four cases (Figure 5.a, Figure 5.b), high- and low-performance groups showed similar relationships between operator interaction factors and work performance factors. As the number of proactive manipulations increased, the alarm reoccurring time increased in both groups (Figure 5.a), and as amount of parameter modification increased, the throughput decreased (Figure 5.b). However, there was a case that shows the opposite tendency between high- and low-performance groups (Figure 5.d). In the case of the high-performance group, the throughput increased as the ratio of reactive manipulation increased, but the low-performance group showed reversed results.

6.2.5 Correlation between Performance Factors. We selected three representative performance factors for different granularity (microscopic to macroscopic) of temporal levels; event-level for trouble reoccurring time, cycle-level for cycle time, and shift-level for throughput. These performance factors were correlated as follows: The Pearson correlation coefficient was 0.45 between cycle time and trouble recurrence time, -0.57 between throughput and trouble alarm recurrence time, and -0.83 between throughput and cycle time. We found that the independent variables affecting each factor differed. For example, although there was a high correlation between cycle time and throughput, the spec change factor affected only the throughput (Table 2). The spec changes events occurred between consequent cycles (e.g., from the end of the previous cycle to the start of the next cycle). Therefore, if spec changes occurred during a shift, the throughput was decreased, but the cycle time was not affected. Furthermore, the operator-HMI interaction factors affecting each level of performance were different. This indicates that some interaction factors are important to observe the microscopic productivity like trouble alarm reoccurring time, however,

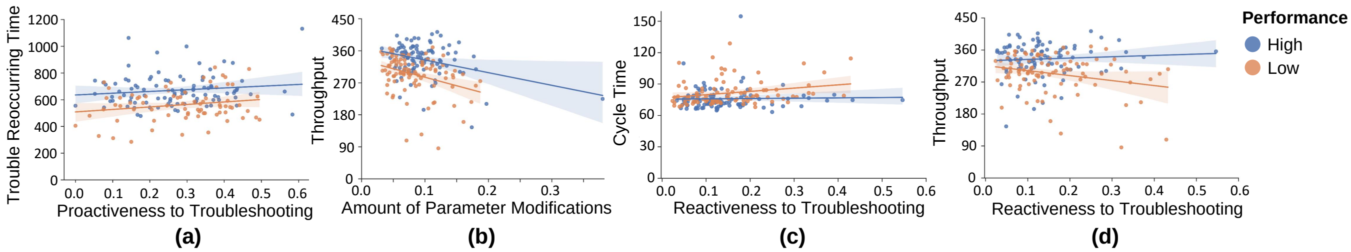

Figure 5: Relationships between three performance factors and significant interaction factors for each. High- and low-performance groups showed similar (a, b) or opposite (d) relationships between operator interaction factors and work performance factors. The light-colored area around line graphs indicates 95% confidence bands

some other factors were important in macroscopic productivity like throughput.

7 DISCUSSION

This study proposed a general data-analysis methodology of data contextualization for behavior modeling by extracting diverse performance-contributing factors from real-world machine data. By integrating data from multiple sources, we could contextualize the machine-generated data and figure out that the work performance varies depending on the operators' behavior patterns. Below, we first discuss the potential of fine-grained, data-driven operator behavior modeling methodology and its generalizability with extensibility. Then we discuss the operator performance variation from modeling results and its utilization as valuable assets to understand their behaviors. Lastly, we discuss potential ethical concerns with machine data-driven operator modeling. Based on our findings, we explore the design implications of context-aware HMI with machine-generated data analytics. The following discussion of our work can contribute to the growing HCI works of human behavior modeling, especially in industrial domains.

7.1 Performance Modeling with Multiple Streams of Machine-Generated Data

7.1.1 Data Contextualization for Fine-Grained Quantitative Analysis. We answered the two research questions by analyzing the operator's behavior and performance based on our methodology. To answer the first research question, we modeled operators' working behavior and performance using real-world machine-generated data by data contextualization. Previous studies have proposed theoretical models that conceptually introduce factors that affect performance such as the operator's individual background [6], external environment [41], and interaction related factors [37, 42]. However, there has been a lack of real-world case studies despite the maturity of the concept of Industry 4.0. Furthermore, existing studies have not attempted fine-grained analysis on machine-generated data at the event, batch, and shift-level that can offer supporting evidence for optimizing human-machine interaction. In this study, we bridged the gap by employing real-world machine-generated data as a new modality and proposing an analysis methodology that enables quantitative measurement of the working behavior, working environment, and performance. Moreover, this approach has the advantage of not burdening factory operators, and researchers can collect data in a non-intrusive manner. As Nasser and Tariz emphasized [45], the critical challenge in the process was integrating and aggregating various data and extracting valid information from relevant data. We collaborated with experts who have worked for a tire company to overcome the challenges and understand the meaning of data from HMI. Reflecting on their domain knowledge, we built metadata to understand how manufacturing contexts are reflected in machine-generated data, which is critical to the proposed data analysis methodology. Based on the metadata, we aggregated the multiple streams of machine-generated data and performed data contextualization by combining the raw data for devising valid factors. Through this study, we demonstrated the potential for data-driven, quantitative operator behavior and performance modeling in the industrial domain.

This research can be applied to other situations that employ the semi-automated manufacturing system and can be extended with applicable contextual factors. For generalizability, we utilized the dataset that follows ISA-95 standards [4], which is a commonly-used international standard to develop automated systems between enterprise and control systems that contain PLC, MES, and HMI data. Similarly, this methodology can be generalized to other semi-automated manufacturing systems that follow ISA-95 standards. For example, it can be applied to manufacturing processes such as semiconductor or automobile assembly, where systems automatically execute processes and require human operators intermittently for troubleshooting. The extracted contexts comprising temporal context, machine state, related events matching, event properties, and working environment are terms that can be derived from a troubleshooting situation in the general manufacturing process. Moreover, we suggest the extensibility of operator behavior and performance modeling by introducing a fundamental methodology that can explain the human-machine collaborative work process. This methodology can be extended to consider other contextual factors such as individual and organizational factors as in existing theoretical frameworks [7, 37, 66]. The extension allows a fine-grained understanding of human-machine interactions.

7.1.2 Performance Variation Depending on Behavior and Environmental Context. Based on our analysis method, we revealed that the performance variation can be explained by the human-machine interaction and work environment factors. The existing theoretical and conceptual models could not explain *how* each factor (e.g., environments and interaction factors) influences performance in a quantifiable manner. We utilized numerical and categorical real-world data to quantitatively analyze the relationship between performance and its related factors. The results clarified which behavioral patterns are important to improve each performance factor. For example, the operator's proactiveness in the human-machine interaction context is important to prevent reoccurring of alarms, while reactiveness is crucial to shortening the production time. The extracted contexts and factors can help to discover phenomena that are difficult to notice without data-driven analysis. In addition, we confirmed that there is a *between-person variation* as well as a *within-person variation* in performance by comparing marginal R^2 and conditional R^2. The fact that the conditional R^2 was higher than the marginal R^2 in all cases showed individual performance differences among operators. In the interview, the experts who have been working for the tire company mentioned that the between-person variation could be caused by considerable individual differences in decision-making between experts with rich experience and novices with less experience.

As such, machine-generated data contains detailed work histories of how operators interacted with machines and which interaction led to better performance. Thus, we claim that the machine-generated data can be employed as a valuable asset that enables us to understand the operator's behavior. The results of this study can be utilized as an objective reference to build an efficient manufacturing process in the industry domain. For example, know-how that had to be grasped subjectively and empirically based on the experience of experts can be objectively acquired and verified through a data-driven approach. Also, in terms of manpower management, managers can use our methodology to schedule working shifts, decide roles, or figure out the working characteristics of each operator. Furthermore, our study showed that the human role is still crucial even in semi-automated systems, where many processes are automated. Although the machine used in our study is a *supervisory control* [19] in which the system is involved in overall processes and humans intervene only when necessary, human behavior was still significant in the performance of the collaboration. Many automated systems, such as driving, manufacturing, home and medical appliances, have supervisory controls that require user intervention in the middle of the automated state. As semi-automated systems are common around us, it is essential to understand fine-grained human behavior in order to optimize the interaction between humans and machines and increase the efficiency of collaboration. When the system frequently requires human intervention, guidelines or standard operating practices for human intervention are necessary to maintain performance such as productivity, quality, and safety at a stable level.

7.2 Implications on the Design of Context-Aware HMI with Machine Data Analytics

Context-specific guidelines are necessary to deal with performance (within- and between-person) variations, so that an appropriate solution can be provided for each type of situation. Guidelines that offer suitable actions can be learned from the accumulated data and automatically delivered to the operators via context-aware feedback through HMI interfaces. As mentioned previously, there is a frequent working sequence when an operator interacts with a machine in a given situation. The mining of interaction sequences can aid in determining the best-performing representative interaction patterns [16]. Furthermore, the operator's behavior can be monitored and tracked in real-time to support the decision-making process. A context-aware HMI will be able to recommend the most appropriate action for troubleshooting. Rich machine data enables the current contexts (machine, material, and human) to be tracked accurately and users to be guided appropriately.

The machine data can enable data-driven personal reflection on behaviors, making it possible to analyze the behavior pattern of an operator and compare it with that of other operators. These data help operators to self-reflect on their performance (e.g., by reviewing poorly performing work sequences). Furthermore, machine data analyses can be used for data-driven personas [40] in manufacturing settings. The concept of the persona, which was presented by Cooper [15] explains the behaviors, goals, and needs of the user [65]. A data-driven persona is an advanced form of persona that incorporates large-scale user data using computational methods [28]. McGinn et al. [40] stated that data-driven personas require contextualized behavioral variables to be used for intelligent interface design.

7.3 Potential Ethical Concerns with Machine Data Driven Productivity Monitoring

Monitoring user behavior in a workspace has been traditionally carried out to improve productivity in manufacturing. In the early 20th century, Taylor attempted to optimize an overall manufacturing process by observing operator behavior in detail and shortening sub-task process times [53]. Toward this goal, Gilbreth pioneered a motion analysis method that aims to reduce unnecessary motion for work efficiency [20]. Recent technological advances in mobile, wearable, and Internet of Things enabled fine-grained behavior monitoring of operators by observing various types of sensor and machine data, which brought forth the concept of *digital Taylorism* [2, 27, 46], a labor management practice involving "new modes of the measurement, standardization, and quantification, decomposition, and surveillance of labor" as in algorithmic and data-driven management [23, 36]. Despite the potential for productivity enhancement, data-driven operator modeling and its use for management may raise potential ethical issues.

First of all, being under surveillance can be stressful for an operator just by its own nature. Surveillance leads to operators' psychosocial risks such as reduced job satisfaction and increased stress and counterproductive work behavior [8]. Furthermore, surveillance will reduce operators' trust in supervisors and management particularly when the purpose, data items, and period of data collection are not transparently shared with operators [51]. In this work, we highlighted that analyzing machine-generated data enables fine-grained user behavior and performance modeling, as well as data-driven behavior optimization (or behavior intervention). Machine data stores all interactions and work histories between operators and machines that could not be observed using conventional wearables or cameras. The data facilitates a detailed understanding of performance, human errors, and work-related behavioral characteristics.

Further, data-driven behavior optimization is feasible by systematically intervening user behavior for performance improvement (e.g., by preempting behavioral violations [61]). This kind of digital Taylorism will pervade semi-autonomous system design, and it may negatively influence how operators perceive their role in the manufacturing process. Behavior quantification and intervention by the intelligent system will cause operators to view themselves as mechanical components of a large, complex, intelligent machine. When working *within* such intelligent machines, operators may experience losing a sense of control in this dehumanized workplace, negatively affecting employees' feelings of autonomy on the job, self-management, and motivation in the long term.

In the new era of digital Taylorism where machine data, as well as other types of monitoring technologies (including surveillance cameras and wearable trackers), are used for behavior quantification and intervention, employers should be aware of trade-offs between productivity monitoring and psychological well-being by carefully investigating ethical and privacy concerns [1]. It is of the utmost importance to ensure that employers must establish the legality of any data to be collected on monitoring targets such as data necessity and proportionality aligned with legitimate business interest [8]. Regarding data privacy, when conducting in-house data analyses, it is recommended to properly pseudonymize machine data. Additionally, there should be safeguards such as individual review of data use and organizational measures for data protection [63]. Redesigning collaborative work between humans and automated systems can possibly mitigate the negative side effect of dehumanizing algorithmic management. We can consider how human intelligence can assist, train, and supervise automation as a joint cognitive system [60] where operators use the data to resolve complex errors and engage in jointly devising strategies for performance improvement. This requires a set of empowering tools that transparently explain decision-making processes, facilitate operator training, and allow customizable machine control [2, 31, 36, 60].

7.4 Limitations

Analyzing machine-generated data with data contextualization provides an opportunity for a fine-grained understanding of the relationship between operators' behavior and work performance. However, it is difficult to capture all manufacturing contexts with machine-generated data. For example, we could not observe operator behaviors when manipulating the machine or materials without HMI usage. Therefore, our methodology is limited in that we cannot capture such operator behaviors that affect manufacturing performance but are unmediated by HMI. Using other data sources may allow researchers to understand operators' behaviors more comprehensively. For example, wearable devices or video-recording using cameras can be used to monitor operators' working behaviors [39, 47]. In our work, we did not consider individual operators' characteristics, such as work experience, age, and level of competence, which can affect work performance. Instead, we focused on analyzing general interaction tendencies from machine-generated data, because we were more interested in how operators' work behaviors are associated with their work performance than the individual operator's characteristics.

8 CONCLUSIONS

We proposed a methodology for modeling human-machine interaction and work performance using semi-automated machine data in manufacturing contexts. We used machine data collected from a tire building machine in a large tire manufacturing company as a case study. We demonstrated that various factors that contribute to work performance can be quantified by integrating and contextualizing machine data from various sources. We defined and extracted the operator-machine interaction, environmental, and performance factors from the data. Our multilevel linear regression modeling revealed that both the interaction and environmental factors influenced the performance significantly. Our methodology and modeling results highlight the possibility of using machine-generated data as sensor data, which enables operator behaviors and their working environments to be understood. These findings suggest new research directions for exploring machine data, such as mining representative interaction sequences, context-aware HMI, data-driven personal training, and developing data-driven personas. We expect that our methodology can be extended to other similar semi-automated manufacturing processes; thus, it serves as a new foundation for the construction of context-aware manufacturing systems using machine data.

ACKNOWLEDGMENTS

This work was supported in part by Hankook Tire & Technology and the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Korean government (MSIT) (No. 2021M3A9E4080780). During the planning and development stages of this study, Chanhee Lee and Jin Uk Shin of Hankook Tire & Technology provided us with invaluable and constructive suggestions. Mr. Lee's expertise as a lead facility engineer for tire building machines was important in our ability to comprehend the tire manufacturing machine's operations and interpret the raw data and operator interactions. Additionally, as a research advisor, Mr. Shin assisted us in determining the overall research directions. We would like to deeply thank Director Jaehee Kim for inspiring our research on human-machine collaboration in semi-autonomous machines.

REFERENCES

[1] Ifeoma Ajunwa. 2018. Algorithms at work: productivity monitoring applications and wearable technology as the new data-centric research agenda for employment and labor law. . *Louis ULJ* 63 (2018), 21.
[2] Moritz Altenried. 2020. The Platform as Factory: Crowdwork and the Hidden Labour Behind Artificial Intelligence. *Capital & Class* 44, 2 (2020), 145–158.
[3] Tayfur Altiok. 2012. *Performance analysis of manufacturing systems.* Springer Science & Business Media, Piscataway, NJ, USA.
[4] ANSI/ISA-95.00.01-2010 2010. *Enterprise Control System Integration - pt. 1.* Standard. International Society of Automation.
[5] Phillip Backus, Mani Janakiram, Shahin Mowzoon, C Runger, and Amit Bhargava. 2006. Factory cycle-time prediction with a data-mining approach. *IEEE Transactions on Semiconductor Manufacturing* 19, 2 (2006), 252–258.
[6] L. Bainbridge. 1983. Ironies of Automation. In *Analysis, Design and Evaluation of Man–Machine Systems.* Pergamon, 129–135.
[7] Tim S Baines, R Asch, Linda Hadfield, JP Mason, Sarah Fletcher, and John M Kay. 2005. Towards a theoretical framework for human performance modelling within manufacturing systems design. *Simulation modelling practice and theory* 13, 6 (2005), 486–504.
[8] Kirstie Ball. 2021. Electronic Monitoring and Surveillance in the Workplace. *Publications Office of the European Union, Luxembourg, doi* 10 (2021), 5137.
[9] Nikola Banovic, Tofi Buzali, Fanny Chevalier, Jennifer Mankoff, and Anind K. Dey. 2016. Modeling and Understanding Human Routine Behavior. In *Proceedings*

of the 2016 CHI Conference on Human Factors in Computing Systems (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 248–260.

[10] Nikola Banovic, Anqi Wang, Yanfeng Jin, Christie Chang, Julian Ramos, Anind Dey, and Jennifer Mankoff. 2017. Leveraging Human Routine Models to Detect and Generate Human Behaviors. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6683–6694.

[11] Heiner Bubb. 2005. Human reliability: A key to improved quality in manufacturing. *Human Factors and Ergonomics in Manufacturing & Service Industries* 15, 4 (2005), 353–368.

[12] A.J. Bunting and A.J. Belyavin. 1999. Modelling human performance in semi-automated systems. In *1999 International Conference on Human Interfaces in Control Rooms, Cockpits and Command Centres*. Institution of Engineering and Technology, Bath, UK, 21–25.

[13] Stephen M Casner, Edwin L Hutchins, and Don Norman. 2016. The challenges of partially automated driving. *Commun. ACM* 59, 5 (2016), 70–77.

[14] Andrea Cherubini, Robin Passama, André Crosnier, Antoine Lasnier, and Philippe Fraisse. 2016. Collaborative manufacturing with physical human–robot interaction. *Robotics and Computer-Integrated Manufacturing* 40 (2016), 1–13.

[15] Alan Cooper. 1999. *The Inmates are Running the Asylum*. Vieweg+Teubner Verlag, Wiesbaden, 17–17.

[16] Alican Dogan and Derya Birant. 2021. Machine learning and data mining in manufacturing. *Expert Systems with Applications* 166 (2021), 114060.

[17] Paul Dourish. 2004. What we talk about when we talk about context. *Personal and ubiquitous computing* 8, 1 (2004), 19–30.

[18] Mica R Endsley. 2017. From here to autonomy: lessons learned from human–automation research. *Human factors* 59, 1 (2017), 5–27.

[19] Mica R Endsley, Betty Bolté, and Debra G Jones. 2003. *Designing for situation awareness: An approach to user-centered design*. CRC press, London, UK.

[20] Frank Bunker Gilbreth and Robert Thurston Kent. 1911. *Motion study: A method for increasing the efficiency of the workman*. D. Van Nostrand Company.

[21] Sonika Gogineni, Kai Lindow, Jonas Nickel, and Rainer Stark. 2020. Applying Contextualization for Data-Driven Transformation in Manufacturing. In *Advances in Production Management Systems. Towards Smart and Digital Manufacturing*, Bojan Lalic, Vidosav Majstorovic, Ugljesa Marjanovic, Gregor von Cieminski, and David Romero (Eds.). Springer International Publishing, Cham, 154–161.

[22] David Gotz, Shun Sun, and Nan Cao. 2016. Adaptive Contextualization: Combating Bias During High-Dimensional Visualization and Data Selection. In *Proceedings of the 21st International Conference on Intelligent User Interfaces* (Sonoma, California, USA) *(IUI '16)*. Association for Computing Machinery, New York, NY, USA, 85–95.

[23] Kathleen Griesbach, Adam Reich, Luke Elliott-Negri, and Ruth Milkman. 2019. Algorithmic control in platform food delivery work. *Socius* 5 (2019), 2378023119870041.

[24] Scott S. Grigsby. 2018. Artificial Intelligence for Advanced Human-Machine Symbiosis. In *Augmented Cognition: Intelligent Technologies*, Dylan D. Schmorrow and Cali M. Fidopiastis (Eds.). Springer International Publishing, Cham, 255–266.

[25] Chunshi Guo, Chouki Sentouh, Jean-Baptiste Haué, and Jean-Christophe Popieul. 2019. Driver–vehicle cooperation: a hierarchical cooperative control architecture for automated driving systems. *Cognition, Technology & Work* 21, 4 (2019), 657–670.

[26] Marie Hoarau, Camilo Charron, and Franck Mars. 2014. Activity Analysis of Expert and Novice Operators in a Semi-Automated Manufacturing Process. In *Proceedings of the 2014 European Conference on Cognitive Ergonomics* (Vienna, Austria) *(ECCE '14)*. Association for Computing Machinery, New York, NY, USA, Article 19, 4 pages.

[27] John Howard. 2022. Algorithms and the future of work. *American Journal of Industrial Medicine* 65, 12 (2022), 943–952.

[28] Bernard Jansen, Joni Salminen, Soon-gyo Jung, and Kathleen Guan. 2021. Data-driven personas. *Synthesis Lectures on Human-Centered Informatics* 14, 1 (2021), i–317.

[29] Christian P Janssen, Stella F Donker, Duncan P Brumby, and Andrew L Kun. 2019. History and future of human-automation interaction. *International Journal of Human-Computer Studies* 131 (2019), 99–107.

[30] Andrew KS Jardine, Daming Lin, and Dragan Banjevic. 2006. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical systems and signal processing* 20, 7 (2006), 1483–1510.

[31] Aniket Kittur, Jeffrey V Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1301–1318.

[32] Naveen Kumar and Seul Chan Lee. 2022. Human-machine interface in smart factory: A systematic literature review. *Technological Forecasting and Social Change* 174 (2022), 121284.

[33] Marcel Langer and Dirk Söffker. 2011. Human guidance and supervision of a manufacturing system for semi-automated production. In *2011 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*. IEEE, Amman, Jordan, 1–6.

[34] Heiner Lasi, Peter Fettke, Hans-Georg Kemper, Thomas Feld, and Michael Hoffmann. 2014. Industry 4.0. *Business & information systems engineering* 6, 4 (2014), 239–242.

[35] Jay Lee, Hung-An Kao, and Shanhu Yang. 2014. Service innovation and smart analytics for industry 4.0 and big data environment. *Procedia cirp* 16 (2014), 3–8.

[36] Min Kyung Lee, Daniel Kusbit, Evan Metsky, and Laura Dabbish. 2015. Working with machines: The impact of algorithmic and data-driven management on human workers. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 1603–1612.

[37] Jue Li, Minghui Yu, and Hongwei Wang. 2018. A taxonomy of performance shaping factors for shield tunnel construction. *Engineering, Construction and Architectural Management* 25 (2018), 574–596.

[38] Yuan Luo, Ya Xiao, Long Cheng, Guojun Peng, and Danfeng Yao. 2021. Deep learning-based anomaly detection in cyber-physical systems: Progress and opportunities. *ACM Computing Surveys (CSUR)* 54, 5 (2021), 1–36.

[39] Takuya Maekawa, Daisuke Nakai, Kazuya Ohara, and Yasuo Namioka. 2016. Toward Practical Factory Activity Recognition: Unsupervised Understanding of Repetitive Assembly Work in a Factory. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (Heidelberg, Germany) *(UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 1088–1099.

[40] Jennifer (Jen) McGinn and Nalini Kotamraju. 2008. Data-Driven Persona Development. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1521–1524.

[41] Dwight P Miller. 1987. Human error and human reliability. *Handbook of human factors* (1987).

[42] Friedhelm Nachreiner, Peter Nickel, and Inga Meyer. 2006. Human factors in process control systems: The design of human–machine interfaces. *Safety Science* 44, 1 (2006), 5–26.

[43] Kevin Nagorny, Pedro Lima-Monteiro, Jose Barata, and Armando Walter Colombo. 2017. Big data analysis in smart manufacturing: A review. *International Journal of Communications, Network and System Sciences* 10, 3 (2017), 31–58.

[44] Shinichi Nakagawa and Holger Schielzeth. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods in ecology and evolution* 4, 2 (2013), 133–142.

[45] T Nasser and RS Tariq. 2015. Big data challenges. *J Comput Eng Inf Technol 4: 3. doi: http://dx. doi. org/10.4172/2324* 9307, 2 (2015), 1–10.

[46] José Luis Vázquez-García María Purificación. [n. d.]. From taylorism to neo-taylorism: a 100 year journey in human resource management. ([n. d.]).

[47] Xia Qingxin, Atsushi Wada, Joseph Korpela, Takuya Maekawa, and Yasuo Namioka. 2019. Unsupervised factory activity recognition with wearable sensors using process instruction information. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 1–23.

[48] Ricardo J. Rabelo, David Romero, and Saulo Popov Zambiasi. 2018. Softbots Supporting the Operator 4.0 at Smart Factory Environments. In *Advances in Production Management Systems. Smart Manufacturing for Industry 4.0*, Ilkyeong Moon, Gyu M. Lee, Jinwoo Park, Dimitris Kiritsis, and Gregor von Cieminski (Eds.). Springer International Publishing, Cham, 456–464.

[49] David Romero, Peter Bernus, Ovidiu Noran, Johan Stahre, and Åsa Fast-Berglund. 2016. The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation Towards Human-Automation Symbiosis Work Systems. In *Advances in Production Management Systems. Initiatives for a Sustainable World*, Irenilza Nääs, Oduvaldo Vendrametto, João Mendes Reis, Rodrigo Franco Gonçalves, Márcia Terra Silva, Gregor von Cieminski, and Dimitris Kiritsis (Eds.). Springer International Publishing, Cham, 677–686.

[50] Priyo Nath Roy, Maniza Armin, S. M. Kamruzzaman, and Md. Emdadul Hoque. 2019. A Supervisory Control of Home Appliances using Internet of Things. In *2019 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, Cox'sBazar, Bangladesh, 1–6.

[51] Jeffrey M Stanton. 2000. Reactions to employee performance monitoring: Framework, review, and research directions. *Human Performance* 13, 1 (2000), 85–113.

[52] Ljiljana Stojanovic, Marko Dinic, Nenad Stojanovic, and Aleksandar Stojadinovic. 2016. Big-data-driven anomaly detection in industry (4.0): An approach and a case study. In *2016 IEEE International Conference on Big Data (Big Data)*. IEEE, Washington, DC, USA, 1647–1652.

[53] Frederick Winslow Taylor. 2004. *Scientific management*. Routledge.

[54] N Viswanadham and Y Narahari. 2015. *Performance modeling of automated systems*. PHI Learning Pvt. Ltd., Delhi, India.

[55] Dakuo Wang, Elizabeth Churchill, Pattie Maes, Xiangmin Fan, Ben Shneiderman, Yuanchun Shi, and Qianying Wang. 2020. From Human-Human Collaboration to Human-AI Collaboration: Designing AI Systems That Can Work Together with People. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–6.

[56] Gang Wang, Xinyi Zhang, Shiliang Tang, Haitao Zheng, and Ben Y. Zhao. 2016. Unsupervised Clickstream Clustering for User Behavior Analysis. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 225–236.

[57] Jinjiang Wang, Yulin Ma, Laibin Zhang, Robert X Gao, and Dazhong Wu. 2018. Deep learning for smart manufacturing: Methods and applications. *Journal of manufacturing systems* 48 (2018), 144–156.

[58] Lindsay R Wilkinson, Kenneth F Ferraro, and Blakelee R Kemp. 2017. Contextualization of survey data: What do we gain and does it matter? *Research in human development* 14, 3 (2017), 234–252.

[59] H James Wilson and Paul R Daugherty. 2018. Collaborative intelligence: Humans and AI are joining forces. *Harvard Business Review* 96, 4 (2018), 114–123.

[60] David D. Woods. 1985. Cognitive Technologies: The Design of Joint Human-Machine Cognitive Systems. *AI Magazine* 6, 4 (1985), 86–92.

[61] Karen Yeung. 2018. Algorithmic regulation: A critical interrogation. *Regulation & Governance* 12, 4 (2018), 505–523.

[62] Shen Yin and Okyay Kaynak. 2015. Big data for modern industry: challenges and trends [point of view]. *Proc. IEEE* 103, 2 (2015), 143–146.

[63] Sven Zenker, Daniel Strech, Kristina Ihrig, Roland Jahns, Gabriele Müller, Christoph Schickhardt, Georg Schmidt, Ronald Speer, Eva Winkler, Sebastian Graf von Kielmansegg, et al. 2022. Data protection-compliant broad consent for secondary use of health care data and human biosamples for (bio) medical research: towards a new German national standard. *Journal of Biomedical Informatics* 131 (2022), 104096.

[64] Mu Zhang, Chien-Ying Chen, Bin-Chou Kao, Yassine Qamsane, Yuru Shao, Yikai Lin, Elaine Shi, Sibin Mohan, Kira Barton, James Moyne, and Z. Morley Mao. 2019. Towards Automated Safety Vetting of PLC Code in Real-World Plants. In *2019 IEEE Symposium on Security and Privacy (SP)*. IEEE, San Francisco, CA, USA, 522–538.

[65] Xiang Zhang, Hans-Frederick Brown, and Anil Shankar. 2016. Data-Driven Personas: Constructing Archetypal Users with Clickstreams and User Telemetry. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5350–5359.

[66] Ji Zhou, Yanhong Zhou, Baicun Wang, and Jiyuan Zang. 2019. Human–cyber–physical systems (HCPSs) in the context of new-generation intelligent manufacturing. *Engineering* 5, 4 (2019), 624–636.