Winter 12-2019

Countering the Fear of Black-boxed AI in Maintenance: Towards a Smart Colleague

Jonas Wanner
University of Würzburg, jonas.wanner@uni-wuerzburg.de

Lukas-Valentin Herm
University of Würzburg, lukas-valentin.herm@uni-wuerzburg.de

Christian Janiesch
Technische Universität Dresden, christian.janiesch@tu-dresden.de

Countering the Fear of Black-boxed AI in Maintenance: Towards a Smart Colleague

Prototype Demonstration

Jonas Wanner
University of Würzburg
jonas.wanner@uni-wuerzburg.de

Lukas-Valentin Herm
University of Würzburg
lukas-valentin.herm@uni-wuerzburg.de

Christian Janiesch
Technische Universität Dresden
christian.janiesch@tu-dresden.de

Abstract

Digitalization forces improved maintenance in shop-floor systems. Companies have begun to upgrade their existing production lines by equipping them with new machinery or sensors. This enables intelligent tracking and control of manufacturing activities. Simultaneously, the advancement of computing power enables complex analyses including the adaptation of machine learning algorithms to gain new knowledge. However, previous research has revealed that intelligent decision support systems are only applied successfully if they are comprehensible for employees within the factory. Therefore, we have developed a prototype based on a comprehensible set of rules for automated anomaly identification in real-time. We include employee's expert knowledge from the very beginning to establish a sense of participation. This is improved and enhanced by techniques from the fields of process mining and machine learning. Thus, the prototype presents previously unknown error correlations in an understandable and descriptive way combining intelligent anomaly detection by a linked knowledge database system.

Keywords

Smart Factory, Maintenance, Process Mining, Machine Learning, Complex Event Processing.

Introduction

Digitization is advancing and with it the importance of information as a strategic resource. This is particularly evident in modern industry, where maintenance is no longer primarily focused on reactive or relatively blind proactive maintenance approaches. Instead, the advances in digital technology reveal considerable potential for the development of predictive and intelligent maintenance systems and procedures (Lee et al. 2006). In this context, maintenance is an activity in which repairs are carried out at certain intervals to economically extend the operating time of production machines (Gandhi and Ng 2018). Despite the underlying potential, however, modern production facilities can overwhelm the human workforce. This seems to be the case both quantitatively in terms of the sheer amount of data as well as qualitatively due to high complexity of today's shop-floor systems (Antoni and Ellwart 2017). To improve this, machine learning algorithms can be used to process the machinery data to derive the appropriate (re-)actions in time.

However, since the decision making of intelligent algorithms is often difficult to understand and certain inaccuracies remain, the question of trustworthiness and acceptance of those systems among the employees arises. This is a serious problem because solving a maintenance issue usually requires physical actions. Only if the corresponding worker believes in the proposed recommendations of the decision support system, the potential derived from the data will be transferred into action. According to well-founded theories regarding the adaptation of innovations such as Rogers (2010), that is foremost experience, comprehensibility, and the ability to observe.

Motivated by the disadvantages of existing solutions, we take up the challenge of creating a prototypical implantation of an accepted, intelligent (machine) colleague. The prototype creates a real-time capable decision support system for the human worker on site. We use a stepwise approach to face the challenge of user adoption due to its innovative character (Wanner et al. 2019). Thus, to improve acceptance, we include the expert knowledge from employees from the very beginning. For comprehensibility we take on a rule-based processing approach of anomaly detection (Adadi and Berrada 2018). Using a real-time monitoring dashboard, we ensure the ability to observe and understand the underlying process (Lee and See 2004) providing explainable artificial intelligence (XAI).

Our prototype concentrates on scenarios in the field of industrial maintenance where abnormal behavior and corresponding maintenance's action can be detected on the basis of current machine status measured by analyzed sensor data.

The Case of Automated Content Providing

The prototype is called *AutoCoP*, an acronym for *Auto*mated *Co*ntent *P*roviding through smart control systems. Its origin is the digitization of the so-called *technical documentation* of machines: a description of functionalities and operating elements as well as notes on maintenance and troubleshooting, created once by a trained editorial team in a time-consuming process. Typically, a manufacturer tries to simulate the operating behavior of the system in various scenarios in advance to make errors determinable and, thus, addressable by concrete recommendations to fix the respective issue. Due to the increasing complexity of the systems, however, the technical editor can no longer fully predict the number of possible complications, malfunctions, and errors (Kröhn et al. 2016). Thus, a full coverage is not economically feasible or even possible, when a manufacturer first delivers his machine to his customer.

Hence, this is a problem at customer's end. If the machine operator cannot reach a solution in case of an occurring problem with the help of the technical documentation, the machine manufacturer or an expert service employee has to be notified. Typically, the next step is finding a solution through telephone support. However, this is often not sufficient, so that a service technician has to come on site to check and solve the maintenance issue. This requires a lot of time for the appointment coordination, the time to travel, the error localization as well as the repair of the error itself. Even then, it is often hard for the service expert to handle the problem in a timely manner due to the machine's high complexity and today's amount of diagnostics data as well as the infeasibility to carry all sort of spare parts with him. If the machine stands still, high opportunity costs follow, so that it must be the goal of every producer to keep these costs as low as possible.

This is where our research comes in and tries to solve the problem of an inefficient maintenance process due to today's new reality in manufacturing by an intelligent decision support system. The idea is that existing knowledge from the digitized technical documentation at the side of the producing company can be reused and extended in a problem-related manner. Thus, the technical documentation must continue to be updated whenever a previously unknown error occurs. This makes knowledge of errors that have occurred reproducible if the error occurs again. When (known) anomalous machine states are detected, they are combined with context information and expert knowledge from the technical documentation. The result is a visual processing that encapsulates complexity by providing targeted error information and step-by-step instructions for corrective action. With the help of intuitive instructions, employees (especially machine operators) with different experience and qualifications can carry out diagnoses and corrections.

The technical realization of the system is based on the use of extensive machine status data and communication possibilities of modern industrial shop-floor systems. With the help of techniques from the fields of data analytics, real-time processing, and machine learning, an intelligent monitoring of production processes becomes possible. This is combined with a knowledge management system (KMS) to store related know-how to known anomalies. Over time, both a reactive and a proactive logic of the whole system allow problems to become predictable even before they occur. However, the implementation faces various challenges. On the one hand, there is the question of the architecture and the corresponding components to realize an appropriate problem identification with regard to the short information latency required in the production environment. On the other hand, it is necessary to integrate the system into the daily production process, where social aspects become crucial. The acceptance of the employee for the new system, which is based on mechanisms of artificial intelligence, is of particular importance. In the given circumstances, we consider complex event processing (CEP) as promising vehicles to achieve this goal of providing XAI.

Prototype Design and Implementation

Architecture and Technical Selection

The underlying structure of our prototype is based on a lambda architecture, which is frequently used in Big Data environments (Kiran et al. 2015). The lambda architecture is a resource-efficient data handling system combining batch and real-time data processing. In addition, we extend this approach with a knowledge layer to address the digitalization of the technical documentation with its extension and reuse. The result is an effective and cost-efficient procedure for real-time analysis in the Industrial Internet of Things (IIoT). We call it the *knowledge-extended lambda architecture*.

In the following, we present the individual components of the prototype and their interrelations. These implement the principles of an efficient maintenance process of today's shop floor reality in manufacturing as presented in the previous chapter. The components are divided into different layers for the knowledge-extended lambda architecture (cf. Figure 1).

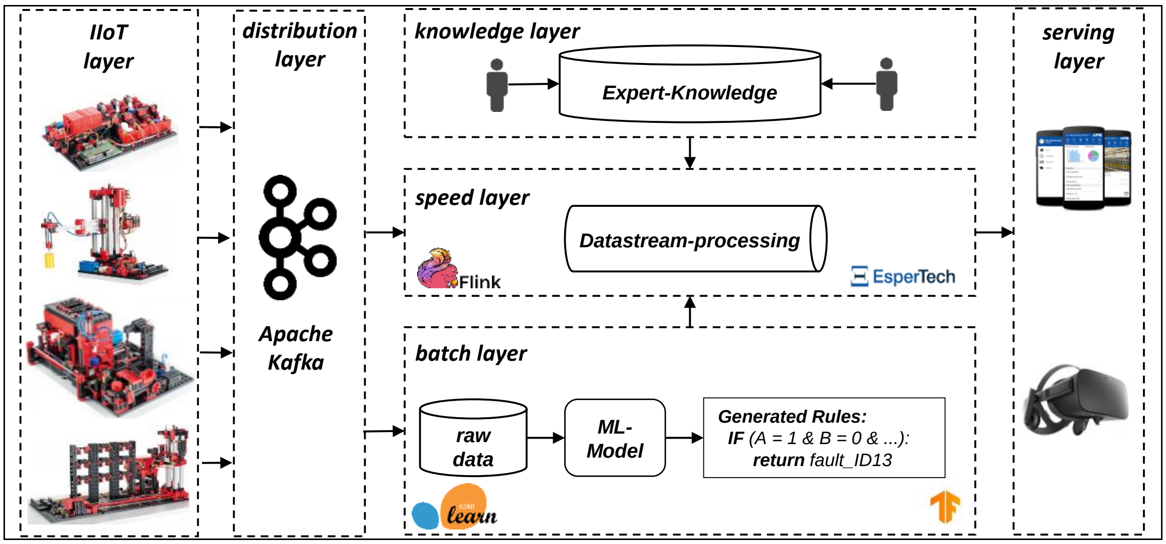

Figure 1. Architecture of the AutoCoP Prototype

IIoT Layer. In the IIoT Layer the data of the shop-floor system is generated and extracted. This is realized through several sensors and actuators, both machine internal and external, to track any changes in the whole production process through its data. In our prototype, the data is generated by a factory simulation model of the model assembly firm *Fischertechnik*, equipped with two industrial *Siemens SPC* SPS control units. The extraction of the data to detect changes comes from more than 200 measuring instances, that is sensors. The type of measurable data contains ambient conditions (temperature and humidity), actuators (pressure and motor voltages), and component status (vibration, speed, gyroscope, and noise).

Distribution Layer. The data retrieved in the IIoT Layer is forwarded into the Distribution Layer which computes several scripts to do so (e.g. *OPC UA* client). Due to the attributes of the data, one can speak of Big Data, characterized by the four V's of variety (different forms of data), volume (scale of data), veracity (uncertainty of data), and velocity (analysis of streaming data). Thus, it is necessary to ensure appropriate implementation of data management (Chen et al. 2012). Publish/ subscribe systems represent a highly scalable and efficient approach to guarantee real-time monitoring. Therefore, these systems are used in the area of Big Data, for example in IIoT. With respect to technological aspects, especially *Apache Kafka* in combination with *Apache Zookeeper* is addressing the necessities of our use case and also many real-world production facilities. These are (1) cache storage to recover previously unread data in case of data processing failures, (2) a subdivision into different topics to support a specific query of certain data items for the individual receivers, (3) implemented redundancies to avoid data loss in case of forwarding failures as well as (4) distribution to dispense the load of the accumulating data to individual clients (Narkhede et al. 2017). In our prototype, we use this approach to provide real-time and fail-safe transmission of sensor data that are distributed across the factory.

Batch Layer. The Batch Layer performs a sequential processing of the received data by the Distribution Layer to learn or improve the detection of anomalous machine behavior. Therefore, the sensor data is evaluated by different machine learning approaches (e.g. decision trees) and converted into a transparent and comprehensible set of rules, to ensure the comprehensibility of the overall system. We use machine learning models from frameworks like *sci-kit learn* or *tensorflow* to do so and translate them by a proprietary middleware developed. This enables diagnostic as well as predictive identification of hitherto unknown problems while continually expanding knowledge about the maintenance process. These rules are made available to the Speed Layer for an immediate monitoring.

Speed Layer. The Speed Layer uses the received rule sets to monitor the incoming stream of sensor data from the Distribution Layer and thus the production process in real-time. CEP technology is used in our prototype in order to do justice to the short information latency in the production environment and to ensure a corresponding traceability and transparency on account of the social aspects. CEP engines are well suited for efficient and highly scalable real-time processing also in combination with machine learning (Wanner et al. 2018). CEP is available from several vendors such as *Apache Flink* or *EsperTech CEP* and automatically adapts the rules for anomaly detection from the Batch Layer. For our prototype we chose *EsperTech CEP* due to its extensive and powerful CEP language *EPL*.

Knowledge Layer. The Knowledge Layer is used to store the expert knowledge of the service technicians needed to fix the failures and can be understood as the digitalized technical documentation layer. Due to the dependency of the knowledge and the detected error, there is a connection to the respective sensor data's error pattern from the Speed Layer. Thus, every error is persisted with an own page in the KMS, containing relevant information about its effective handling. Our prototype is based on a proprietary KMS for the research project, connected to a relational *MySQL* database. In the case of an error occurring in the Speed Layer, the linked knowledge in the KMS is made available to the Serving Layer.

Serving Layer. The Serving Layer operates as the interface for the employees to the AutoCoP system within the factory. Through this interface, employees can use mobile applications, dashboards, and augmented reality glasses to inform themselves about the current status of the production process. In the case of an already known problem occurring in the system, they receive a timely notification with helpful recommendations on what to do. Otherwise, this layer will access the expert knowledge via an input mask and forward it to the KMS to be stored. For the knowledge representation of the real-time observations, we use a proprietary dashboard based on *Angular JS*. For the knowledge representation of the linked technical documentation, we use *Angular JS* as well.

Technical Realization and Illustration

After the setup and connection of the overall system has been implemented, the logic unit must be taught rules to ensure real-time detection of abnormal machine behavior and to be able to link its detection to relevant information on how the resolve the issue. We use a multi-level concept for this purpose, which comprises three steps.

First, local empirical values and knowledge are converted into rules. The shop-floor system operators and internal maintenance technicians are interviewed and their – at times tacit knowledge – is extracted. We also extract information from the technical documentation. Each known anomaly, as a sequence of *n* events and their characteristics, is assigned its own rule and an entry in the KMS to address it. This also ensures the early involvement of employees in the change process and builds experience in the use of the system, as one major step towards innovation adaption. Thereby, it enables the system to detect several errors right from the start and to offer advice.

As a second step and to increase the power of the system, several activities are carried out in parallel. On the one hand, if a new (reactive) error is detected its data pattern is recorded as well as the expert information of its correct handling. The recording of the error pattern also applies to the errors detected by the system to strengthen its reliability. On the other hand, the machine status data is stored in a local database (here *MySQL*). Thus, after some time, process mining techniques and visual data analysis can be used to improve or even extend the existing rules from the CEP rule stack. This makes it even easier to identify abnormal process flows in the future. When no rules or knowledge have yet been developed for this purpose in the system, the to-be process flow of the testing environment can be used to pre-define rules that can be used in the final system.

In order to improve the manual and time-consuming character of step 2, we automated most of the work in step 3. Correspondingly, this requires extensive data, whereby abnormal machine behavior in the data series can be learned with the aid of techniques from the field of machine learning. In order to further ensure comprehensibility and the ability to observe and, thus, to ensure employee acceptance of the decision support system recommendations, the rules identified by algorithm must be converted into (comprehensible) CEP rules. This is done by a middleware we developed, which performs the translation. As an input it requires the name of the (supported) type of machine learning algorithm. The rules improve or even extend the existing CEP rule stack so that, among others, better limits for specific errors are defined automatically. Also, proactive warnings can be issued via appropriate association rules of machine learning algorithms to identify and address maintenance problems ahead of time.

The iterative process of steps 2 and 3 provides an adequate problem coverage over time for maintenance purposes of the machine in use. Table 1 provides some exemplary CEP queries from the rules stack. In order to understand the readability of CEP rules and the most relevant information linked, examples can be found in Table 1, sorted by attributes.

ID	Name	Declaration	KMS_shortEntry	Area
1	TransportE B_to_C – WARNING	`String CEP_1 = "@name('Error TransportE B_to_C detected, ID_1, WARNING') SELECT * FROM pattern [every (a=vacuumPad(compressor=1) AND a=vacuumPad(motor_turn_ counterclockwise=1) -> (timer:interval(12.3 sec) AND NOT b= vacuumPad(oven_lightB=0))];"`	Oven has received no component, error during transfer	C
2	CommunicationE SAW lost - INTERRUPT	`String CEP_2 = "@name('Error CommunicationE SAW lost detected, ID_2, INTERRUPT') SELECT * FROM pattern[every(a=lightBSaw(sensor2037 =1) AND b=actuatorSaw(sensor2030=0) AND c=vibrationSaw(sensor2033<342) .win:time_batch(3 sec)];"`	Saw does not process the component, communication failure	C
n	…	...	…	…

Table 1. Exemplary CEP Queries from the Rule Stack

Figure 2 comprises a conceptual and schematic process diagram in the Business Process Model and Notation (BPMN). It explains the process flow of and the decisions taken by the AutoCoP prototype. The Speed Layer continuously monitors the incoming data stream for anomalies or behavior indicating future anomalies. While there is no anomaly detected, the system shows green lights on its dashboards. When the Speed Layer detects an anomaly, it displays a red light on the corresponding dashboard. It then transfers the complex event to the Knowledge Layer. The Knowledge Layer searches for corresponding patterns in the KMS. If it finds an associated entry in the KMS, it links the pattern of the occurred anomaly to the respective KMS entry. Finally, it displays the knowledge about the anomaly to the user as well as the anomaly's sensor data pattern. If there is no associated entry in the KMS, the Knowledge Layer cannot pass any knowledge to solve the problem to the user. In this case, a user interface in the Knowledge Layer requests new knowledge from the user or a specialized technician remotely or on site. It asks to evaluate the sensor data and to suggest an associated KMS entry or to create a new error documentation (and possible means to address the defect).

Figure 3 demonstrates the real-time monitoring dashboard developed for the prototype, which gives the employee on site the possibility of a live view of machine behavior. If the CEP unit detects an error (reactive) or an imminent problem (proactive), this is color-coded on the corresponding production section in the right corner shop-floor system illustration, and a corresponding notification is displayed. In the case at hand, a (reactive) error has occurred in production section C. It is displayed by the system. The dashboard therefore provides the employee with a direct overview of the nature of the problem (traffic light circuit in the top left-hand corner) as well as a brief overview of the error with the ID, timestamp, involved sensors, a short error description, and the errors' module localization. Clicking on the ID of the error forwards the employee to the linked knowledge entry in the KMS. There he receives an explanation why the recommendation was made by the decision support system (CEP rule and anomaly data pattern) and a step-by-step instruction on how to solve the problem. A commenting function allows to report on his experience. In addition, he can close the maintenance issue by clicking a button or forward it to the next service level.

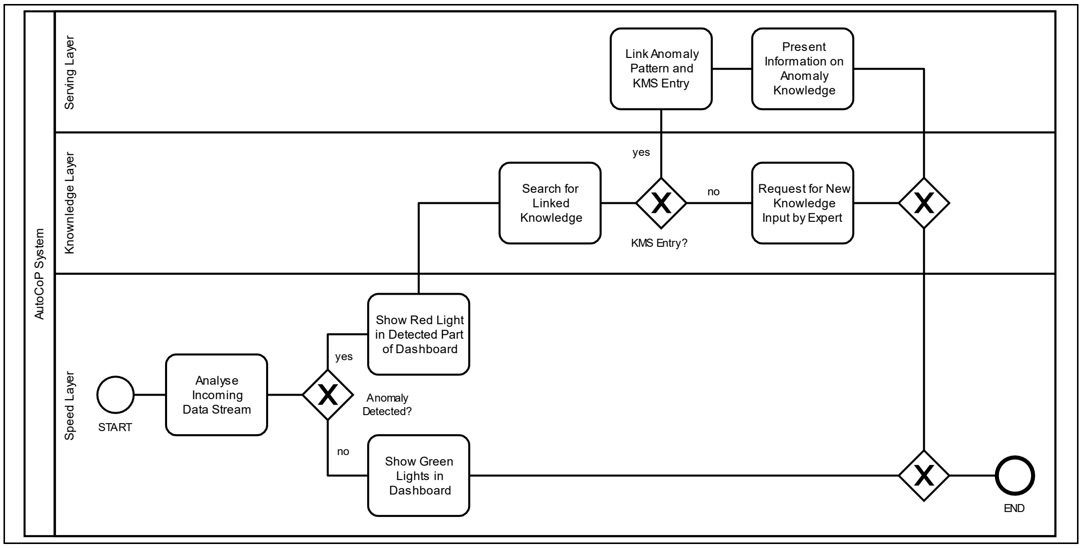

Figure 2: Process Flow of AutoCop Prototype

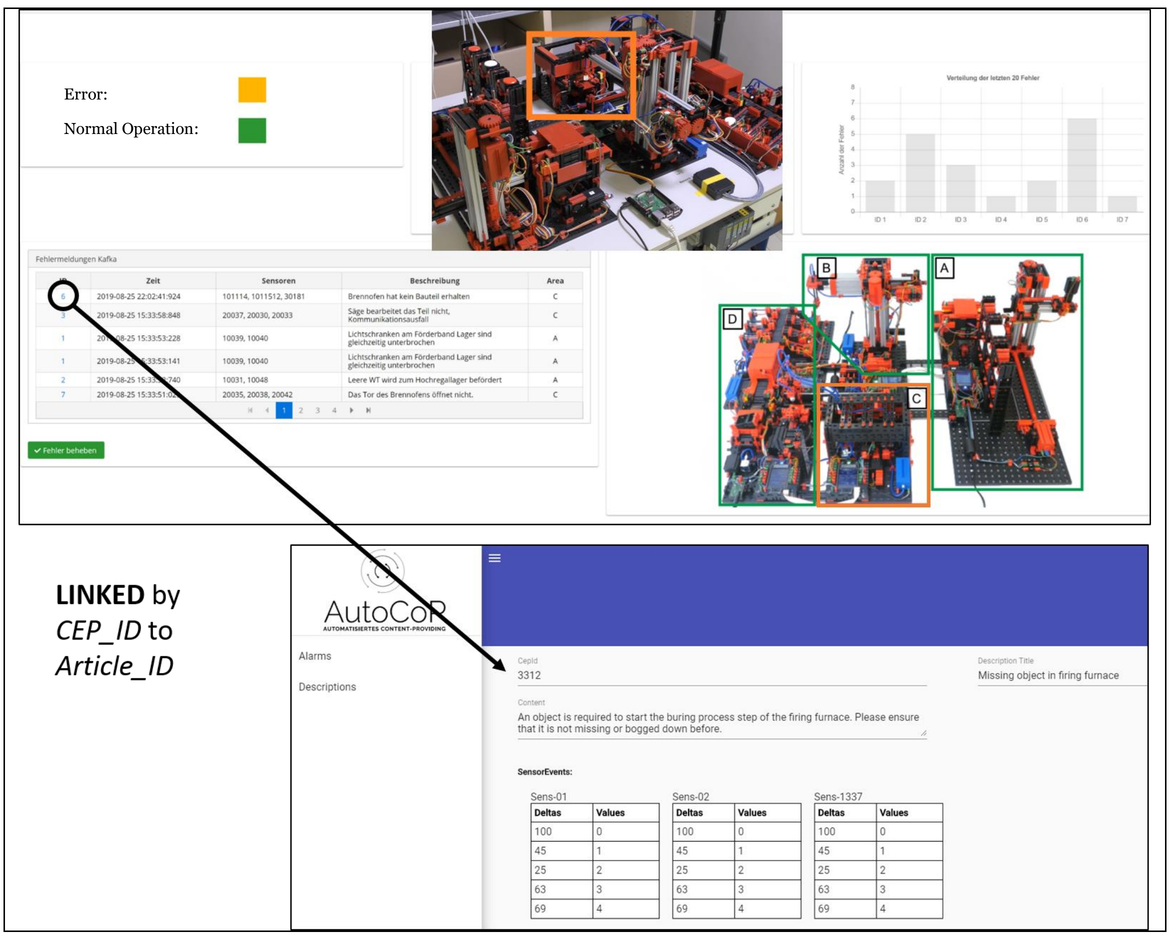

Figure 3: Demonstration of the AutoCoP Prototype

Evaluation and Outlook

The development of the prototype shows that the integration of machine learning approaches can improve the identification of new and complex issues in shop-floors systems. For our *Fischertechnik* model, we were able to automatically generate a prediction of occurring failures across several assemblies. In particular, it was necessary to convert the knowledge extracted from experts and the knowledge gathered through training with machine learning into human-readable rules. We achieved this by using a real-time capable CEP unit, which is connected to a knowledge database to encapsulate the complex data and abstract it into easily understandable information. A real-time decision support system ensures an overall monitoring of the underlying procedure at any time.

We have successfully used the prototype within our research project. As a next step, we plan to evaluate our prototype with a German pipe manufacturer. The contact so far has shown a high degree of willingness and support on behalf of the company, which mirrors the practical need for an assistance system. At the current stage we are primarily using machine learning algorithms such as support vector machine, principal component analysis, or classification trees. In the future, we want to go one step further and explain the results of deep neural nets using an encoder/ decoder model and include them in the prototype. Thereby, we expect to transform the large potentials of complex and almost untraceable deep neural nets into an XAI approach for future use. These results could also be used to enhance the conceptual modeling of process-based monitoring systems (Janiesch and Matzner 2019).

Funding

This research and development project is funded by the Bayerische Staatsministerium für Wirtschaft, Landesentwicklung und Energie (StMWi) within the framework concept “Informations- und Kommunikationstechnik” (grant no. IUK536/003) and managed by the project management agency VDI+VDE Innovation + Technik GmbH.

References

Adadi, A., and Berrada, M. 2018. "Peeking inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI)," IEEE Access (6), pp. 52138-52160.

Antoni, C. H., and Ellwart, T. 2017. "Informationsüberlastung bei digitaler Zusammenarbeit–Ursachen, Folgen und Interventionsmöglichkeiten," Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO) (48:4), pp. 305-315.

Chen, H., Chiang, R. H., and Storey, V. C. 2012. "Business Intelligence and Analytics: From Big Data to Big Impact," MIS Quarterly (36:4), pp. 1165-1188.

Gandhi, K., and Ng, A. H. 2018. "Machine Maintenance Decision Support System: A Systematic Literature Review," Proceedings of the 16th International Conference on Manufacturing Research, Skövde: IOS Press, pp. 349-354.

Janiesch, C. and Matzner, M. 2019. "BAMN: A Modeling Method for Business Activity Monitoring Systems," Journal of Decision Systems, DOI: 10.1080/12460125.2019.1631682, forthcoming.

Kiran, M., Murphy, P., Monga, I., Dugan, J., and Baveja, S. S. 2015. "Lambda Architecture for Cost-Effective Batch and Speed Big Data Processing," 2015 IEEE International Conference on Big Data, Santa Clara, CA: IEEE, pp. 2785-2792.

Kröhn, M., Reuter, N., Hahnle, M., Prochus, T., and Doll, K. 2016. "Gestenbasierte Prozessanalyse," atp magazin (58:12), pp. 26-33.

Lee, J., Ni, J., Djurdjanovic, D., Qiu, H., and Liao, H. 2006. "Intelligent Prognostics Tools and E-Maintenance," Computers in Industry (57:6), pp. 476-489.

Narkhede, N., Shapira, G., and Palino, T. 2017. Kafka: The Definitive Guide: Real-Time Data and Stream Processing at Scale. Sebastopol, CA: O'Reilly.

Rogers, E. M. 2010. Diffusion of Innovations. New York, United States: Simon and Schuster.

Wanner, J., Herm, L.-V., Hartel, D., and Janiesch, C. 2019. "Verwendung Binärer Datenwerte Für Eine Ki-Gestützte Instandhaltung 4.0," HMD Praxis der Wirtschaftsinformatik (56:6), pp. 1-14.

Wanner, J., Wissuchek, C., and Janiesch, C. 2018. "Machine Learning und Complex Event Processing: Effiziente Echtzeitauswertung am Beispiel Smart Factory". Proceedings of the 14. Internationale Tagung Wirtschaftsinformatik (WI), Siegen: AIS, pp. 47-61.