Towards building blocks for predictive analysis of HVAC systems

Nikola Ivanovic[1,*,†], Benjamin Nast[1]

[1]University of Rostock, Albert-Einstein-Str. 22, 18059 Rostock, Germany

Abstract

This paper presents a modular approach for anomaly detection in heating, ventilation, and air conditioning (HVAC) systems by combining domain-specific knowledge and machine learning (ML) applications within reusable building blocks (BB). Based on a real-world industrial case study, which is aligned with current regulatory requirements, the approach supports digital transformation. It illustrates how technical capabilities can be connected with business goals such as energy efficiency, compliance, and service innovation. Machine learning-building blocks (ML-BB), validated through rule-based logic derived from expert knowledge, demonstrate that such components can be reused across use cases with minimal customisation. The feasibility study presented in this work shows that applying the ML-BB can improve the energy efficiency of HVAC systems and reduce the time technicians spend on energy inspections, which are so far done manually. The work contributes to business-IT alignment (BITA) by enabling scalable, domain-driven artificial intelligence (AI) integration into operational workflows.

Keywords

building blocks, machine learning, IoT, HVAC, artificial intelligence

1. Introduction

Heating, ventilation, and air conditioning (HVAC) systems significantly influence indoor air quality and are responsible for approximately 40–60% of energy consumption in buildings [1]. The inefficiency of the HVAC systems has a notable impact on the energy consumption and, therefore, on environmental pollution and energy costs. From this perspective, it is essential to decrease environmental pollution and to protect the end-users by increasing the efficiency of HVAC systems and reducing their energy consumption caused by the inefficiencies in the system performance.

Furthermore, the demand for digitalisation and optimisation comes not only from the environmental and business perspectives but also from regulatory requirements. Since the beginning of 2025, a new regulatory requirement in Germany mandates that HVAC providers integrate systems for monitoring energy consumption. This obligation is detailed in the industrial case study described in section 3.

As most HVAC systems the company under study maintains are not yet adequately digitised, system inspections are conducted manually on-site by technicians. As a result, significant discrepancies between estimated and actual energy efficiency levels are detected. This practice often leads to significant discrepancies between estimated and actual energy efficiency levels.

Through the analysis of inspection reports in a previous research project, the findings indicate that up to 30% energy savings could be achieved through appropriate monitoring practices [2]. These potential savings are further supported by existing studies, which demonstrate that faults can be detected and addressed using low-cost technologies [3, 4]. Consequently, the digitalisation needs of the HVAC systems and the compliance with regulatory frameworks, such as §71a of the Gebäudeenergiegesetz (GEG), are recognised as a key driver of digital transformation.

BIR-WS 2025: BIR 2025 Workshops and Doctoral Consortium, 24th International Conference on Perspectives in Business Informatics Research (BIR 2025), September 17-19, 2025, Riga, Latvia

*Corresponding author.

†These authors contributed equally.

nikola.ivanovic@uni-rostock.de (N. Ivanovic); benjamin.nast@uni-rostock.de (B. Nast)

0009-0006-3874-7216 (N. Ivanovic); 0000-0003-4659-9840 (B. Nast)

To support the digital transformation of the observed company, in a previous project, we developed a system capable of providing performance insights into HVAC operations through the use of Internet of Things (IoT) sensors [2, 5]. In the current project, we build upon this foundation by using machine learning (ML) applications to analyse sensor data, detect system irregularities, and make the system more energy-efficient.

In this work, a collection of analysis building blocks (A-BB) and machine learning building blocks (ML-BB) for intelligent data analysis in HVAC systems is proposed. The proposed building blocks (BB) represent modular ML applications with well-defined interfaces, enabling reuse and minimal customisation. Scalable, domain-driven artificial intelligence (AI) integration into operational workflows is made possible, and thus, it contributes to business-IT alignment (BITA). By applying the ML-BB, energy efficiency can be improved, and the time technicians spend conducting energy inspections can be reduced through task automation.

Since the method for creating BB is not described in this paper, we want to point out that the method is proposed for a separate publication that is currently being prepared. The proposed method describes a systematic approach to designing and deriving BB from a previously created enterprise context model. It makes the ML application structured in modular BB linked to the business processes and data, and enables their reusability across similar enterprise contexts. In this study, we present the ML-BB as outcomes of that method, highlighting their applicability and benefits to the company.

The main contribution of this paper is to demonstrate the practical application of the proposed ML-BB for analysing HVAC data, with a particular focus on identifying anomalous system behaviour and its associated operational costs. The remainder of the paper is organised as follows: Section 2 gives background to BITA and digital transformation, monitoring of HVAC systems, and BB. Section 3 outlines the industrial case study. Section 4 details the application of ML-BB for HVAC system data analysis. Section 5 concludes the paper.

2. Theoretical background

2.1. Business-IT alignment and digital transformation

BITA can be seen as the appropriate and timely application of information technology (IT) in line with the business goals, strategy, and needs [6]. It has a big impact on the performance of organisations [7]. In general, four dimensions can be divided: strategic, structural, social, and cultural, whereby the strategic dimension is researched most [8]. To achieve positive effects, such as increasing efficiency or flexibility in business and IT, all four dimensions need to be considered [9]. As IT is getting increasingly important and has a high impact on efficiency, the importance of BITA is also growing [10].

BITA can be seen as an important factor for digital transformation [10, 11]. In [12], the authors show that it helps enterprises to design and implement organisational flexibility regarding structure, IT infrastructure, and workforce, which is necessary for successful digital transformation. Digital transformation is an important topic in information systems research [13] and for practitioners [14]. It can be described as the adoption of innovative digital technologies, e.g., digital twins, big data, or AI, in the digitalisation of an organisation's business model and its operation [15]. Since it affects not only the products and services enterprises offer, but also how they work [16], an organisation-wide transformation program is needed.

2.2. Monitoring of HVAC systems

Motivated by the reliable climate control, improved indoor air quality, and thermal comfort, many commercial and industrial buildings employ HVAC systems [17]. Even though these systems are increasing the air quality and thermal comfort, they are still accounting for approximately 40-60% of the energy consumption of the buildings [1]. Optimising their energy consumption is inevitable to reduce environmental pollution and protect the end-users from the unneeded energy cost caused by inefficiency in the system operations. In recent years, the latest technologies, such as IoT sensors and

AI, have already found their application in this domain. In [2], the authors applied low-cost sensor technology using a minimally invasive approach to assess HVAC system performance. The collected data enabled the development of various ML and deep learning (DL) applications. ML and DL offered advantages such as early fault detection to identify irregularities in HVAC operations and predictive maintenance to minimise system downtime. ML and DL techniques have been successfully applied to optimise HVAC system operations; however, most studies remain experimental, and only a few have been implemented in real buildings with post-occupancy evaluations [18].

Techniques for diagnosing and detecting errors can essentially be divided into knowledge-based and data-driven [19]. Knowledge-based techniques use existing domain knowledge to develop rules or models for fault detection. They are mostly very complex, require a lot of input from domain experts, and are hard to adapt to different use cases, as they are developed for a specific HVAC system under specific conditions [20]. Data-driven techniques are primarily based on automatically extracting pattern similarities for fault detection, using data-driven approaches [19]. As these require adequate data sets for faulty and non-faulty operations, in most cases, they cannot be used for newly installed systems or changing operating conditions.

2.3. Building blocks as reusable AI components

We performed a literature study investigating existing concepts in structuring and reusing AI applications. For this purpose, we developed a Scopus search query to identify BB, modules, architecture patterns, and other synonyms that serve similar purposes as BB and could be used for structuring and reusing AI solutions. Since BB is a broad term used in many different disciplines, not many papers describe BB or its synonyms in a context relevant to this work. The results were very limited, especially in the context of integrating the technical aspects of AI solutions into business architecture. In this paper, we present a definition of BB that is relevant for our work. The complete literature study and the process by which it was conducted will be published in a separate paper.

Therefore, we envision BB as self-contained, reusable elements [21] that encapsulate business processes, services, and data structures [22], and extend across data, application, and technology layers of the enterprise architecture. They support alignment between business goals and IT capabilities [22, 23], facilitate modular system design, and enable traceable data transformation across the mentioned architecture layers. These elements are characterised by a defined context of the business service, provided and required interfaces between the components, and their relationships.

Finally, we consider the Input-Processing-Output (IPO) model as a suitable structure for both the A-BB and ML-BB. It is a well-established architecture pattern in research for structuring system behaviour in software engineering and systems analysis [24]. A process is described, in which a system receives data (Input), applies a transformation (Process), and produces a result (Output). In this way, we want to support the design's modularity and transparency to increase the BB's reusability.

3. Industrial case study

The work described in this paper is part of an ongoing research project in collaboration with a small and medium-sized enterprise specialising in the construction, operation, and maintenance of HVAC systems. A significant problem in operating these systems is the complexity of their interacting components, and that they are designed individually for each customer. The increasing automation of control technology in industrial facilities and public buildings has led to a rapid growth in the amount of data being generated. Therefore, intelligent data processing is necessary to execute associated processes efficiently.

In accordance with §71a of the GEG in Germany, there is a new legal requirement for non-residential buildings with a heating or cooling system capacity above 290kW. These buildings are required to implement a building automation and control system by December 31, 2024. This system must include digital energy monitoring that enables continuous energy consumption tracking and analysis across all relevant energy sources and building systems. In addition, the system must provide access to the data via open and configurable interfaces, allowing for vendor-independent evaluation [25].

We developed an IoT-based diagnostic tool to facilitate comprehension of the dynamics within the systems and form a basis for new types of business services. Sensors are installed to get a minimal set of measurement data, which is expanded with appropriate physical calculations, to gain deeper insights into the behaviour of a system. To expand our data-driven analysis with domain knowledge, we add a standard set of information about the components of each system (configuration data), which can be found in the device specifications provided by the manufacturers, and in parts are measured under certain conditions in the system. This helps to classify the systems and determine which calculations can be applied. The implemented solution allows for diagnostic support for potential improvements within the systems and the operating processes of the case study company. It can process a large amount of data from different sources and integrate seamlessly into the company's daily operations. More details on the infrastructure and practical applications can be found in our previous work [2, 5].

4. Feasibility study

4.1. Analysis-building block

For the feasibility study, we decided to select the available analysis building block (A-BB) shown in Fig. 1, which describes domain-specific knowledge on how the inspection reports are conducted for the detection of the dehumidification anomaly check service. Furthermore, we selected the suitable ML-BB shown in Fig. 3 to automate the anomaly detection for the same business service using already existing ML applications.

Starting with the A-BB shown in Fig. 1, we observe that the BB follows an IPO structure. On the left side of Fig. 1, the business processes responsible for collecting data required for further analysis are depicted. These include sensor data and system configuration data, which describe the components of the HVAC system. In addition, use case-specific information, such as threshold values, is described. Based on this information, the physical calculation process is executed. In this step, raw sensor data is extended to derive additional insights. The resulting output is represented by the business object element *calculated data*.

In order to detect the anomalies and get insights into whether the components of the HVAC system are functioning as intended, the anomaly checks are performed on the calculated data in the following step. As shown in Fig. 1, the data object *analysed data* indicates that the `check-dehum1` returns `True`. This is the result of the anomaly check process, which means that the dehumidification component of the system is not performing correctly. In the final part of the A-BB, a report is generated, in which all calculated and analysed data are visualised and described. The required types of visualisations are also specified using business object elements, resulting from the business objectives.

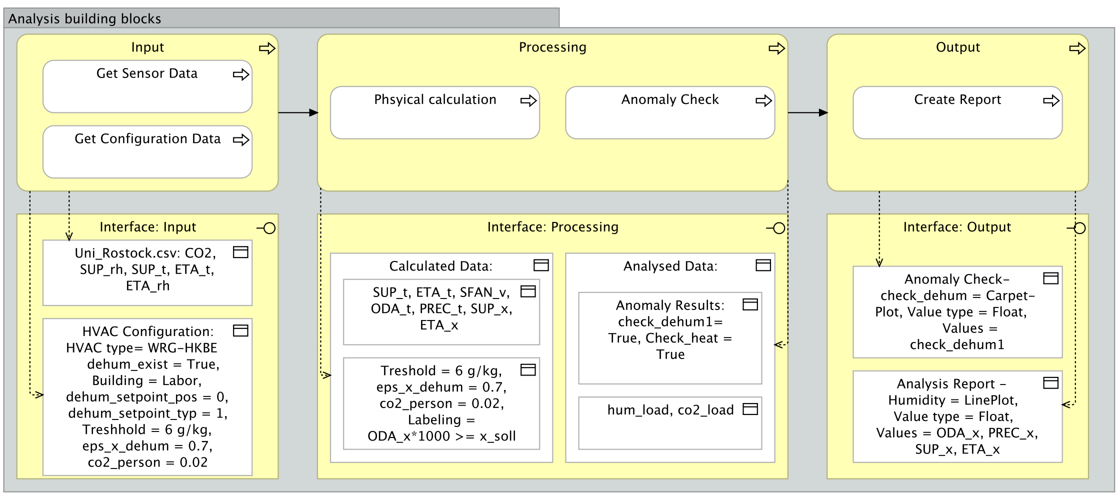

Figure 1: Analysis-building blocks (A-BB).

4.2. Machine learning-building block

Based on the calculated and analysed data of the A-BB, we identified the suitable ML-BB shown in Fig. 2. The ML-BB follows the same IPO model structure as the A-BB but describes the processes and data required or provided by the ML application. Furthermore, it outlines how the data is transformed throughout the ML pipeline. As shown on the left side, the processes that define how the data is accessed from the respective databases are described.

From the A-BB, we know which data should be accessed and transformed for the ML-based anomaly detection in the subsequent processing step. In the next step, the data is processed through the ML pipeline. Here, the data is cleaned, features are generated, and the ML model is trained and tested. In the Output part of the ML-BB, the processes and data needed to provide results of ML anomaly detection and their interfaces are visualised. The practical results of the ML-BB are outlined below.

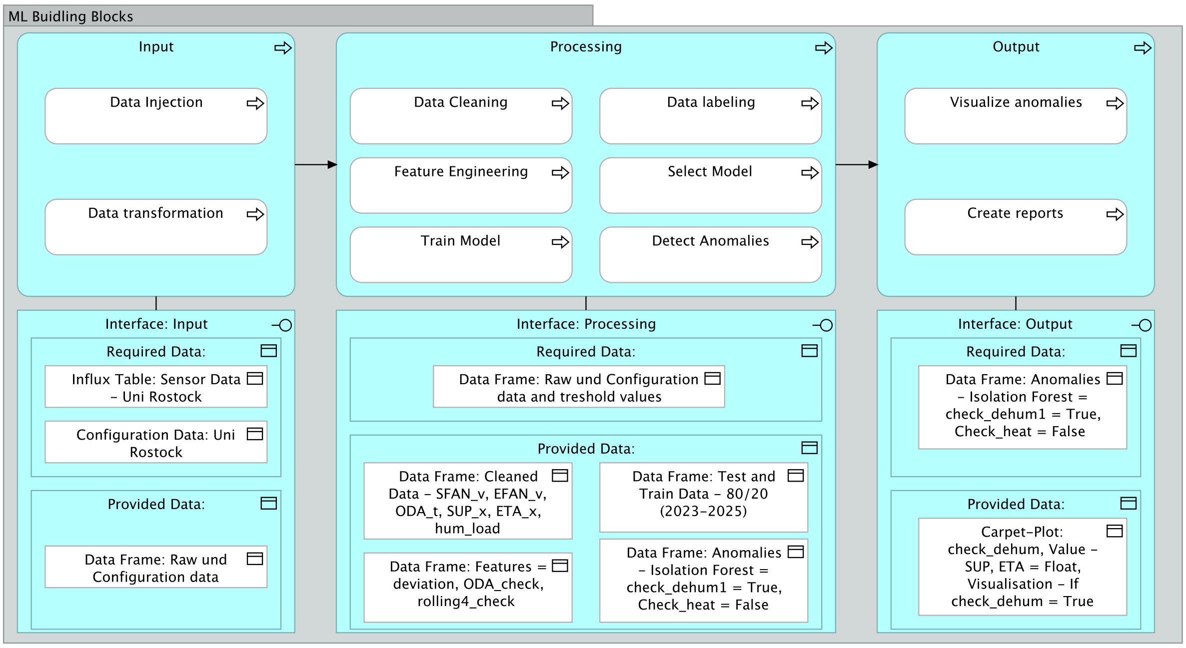

Figure 2: Machine learning-building blocks (ML-BB).

4.3. Results of ML anomaly detection

As previously described, we incorporated the ML-BB shown in Fig. 2 to support the detection of dehumidification anomalies in the HVAC system. The selected ML-BB outlines the components of the ML pipeline that are used to detect dehumidification anomalies using the Isolation Forest algorithm. The ML-BB was selected based on the Input and Output data of the corresponding A-BB and, therefore, ensures compatibility with the business objectives. Since the HVAC data we are working with is not labelled, comes in large volumes, and has significant data noise and outliers, we decided to use an ML model that fits those conditions. We chose the Isolation Forest model because it does not need labelled data, handles large datasets well, and is solid when dealing with outliers and noisy data.

Before presenting the results, we must highlight that we applied a rule-based logic to define domain-specific anomaly conditions based on expert knowledge. The rule-based classification helps us validate the anomalies the ML model detects and increases the overall reliability and trust in the results of the ML-BB. Moving to the results, we apply the ML-BB illustrated in Fig. 3 to detect whether the HVAC system is functioning correctly. The anomalous condition is defined as a situation in which the outdoor humidity rises above a use case-specific threshold, and the HVAC system fails to activate the dehumidification function to reduce the indoor humidity below that threshold.

We use the three features defined in the ML-BB shown in Fig. 2 to detect such conditions. The first feature, `deviation`, measures the difference between the actual and target humidity. The second feature, `ODA_check`, captures the difference between outdoor and target humidity. The third feature, `rolling4_check`, analyses how often dehumidification is activated within a selected range of the data. This feature was helpful to filter out false positives that usually appear during short-term transitions, like when the system turns on or off. We only consider anomalies that are occurring at least four consecutive time steps to make the detection more reliable and less sensitive to noise occurring during transitions. Temporal checks help detect irregularities rather than brief fluctuations, which are common in HVAC systems.

In the next step, we interviewed HVAC domain experts to gain the necessary knowledge for detecting dehumidification anomalies. Based on their input, we implemented a set of rule-based conditions and used them as a validation mechanism for the anomalies detected by the ML model. To make validation of anomalies easier, we then visualised both rule-based and ML model-detected anomalies within the same chart in Fig. 3. The rule-based anomalies are visualised as purple dots and labelled as *Rule-Based Anomaly*, while the anomalies detected by the ML model are marked as red crosses in the same chart and labelled as *ML Detected Anomaly*.

To ensure that the ML model was aligned with expert-defined anomaly rules, we trained the ML model with the data points flagged by the rule-based method. In the next step, the ML model was applied to the whole dataset to predict anomalies across the entire dataset. The results were visualised using Plotly, showing the rule-based and ML-detected anomalies alongside actual humidity, threshold humidity, and outdoor humidity.

The initial findings show that the Isolation Forest model successfully captures most of the anomalies identified by the rule-based method. These findings confirm the validity of the ML-BB-detected anomalies and suggest that the ML-BB can be reused in similar use cases.

Lastly, the defined ML-BB can help improve the energy efficiency of HVAC systems by detecting these irregularities. They also reduce the manual work needed from technicians, since energy inspections no longer have to be done by hand, as the energy inspection process is conducted automatically with the help of ML applications contained in the ML-BB.

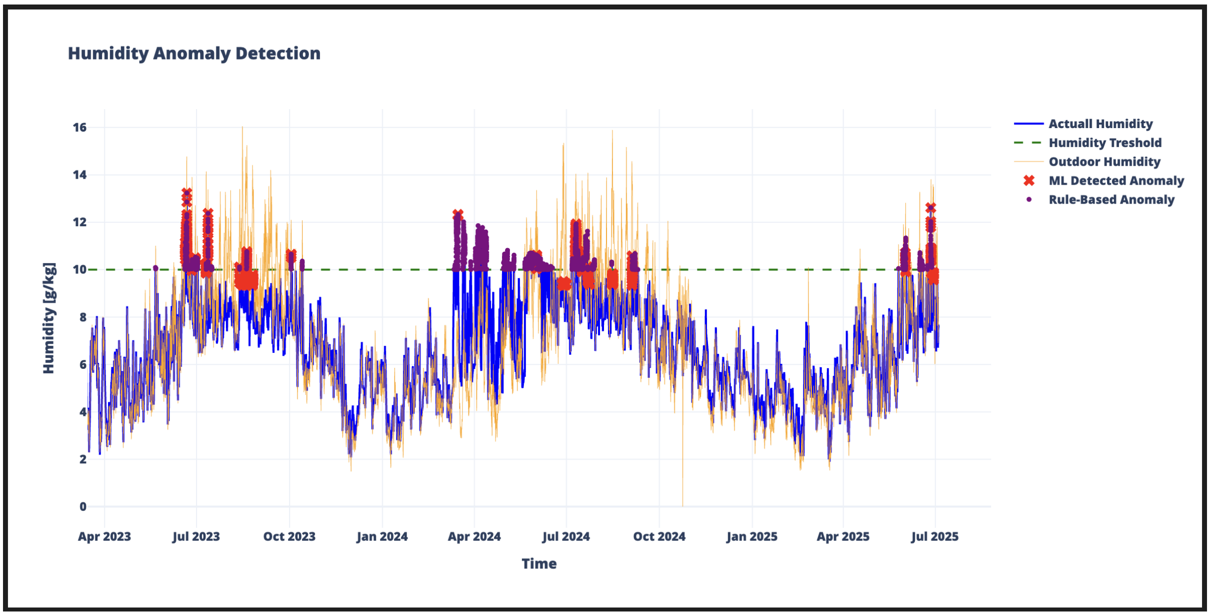

Figure 3: Dehumidification anomaly detection.

5. Conclusion

Using IoT sensors to monitor HVAC systems offers great potential for improving energy efficiency. These systems are often insufficiently digitised and account for significant energy consumption in buildings. Based on a real-world industrial case study aligned with regulatory requirements such as the GEG, the proposed approach demonstrates how BB that integrate domain-specific knowledge with ML applications can be effectively combined to detect dehumidification-related anomalies in HVAC systems. In that way, the structure and logic of the BB outline the alignment between the technical capabilities of the ML model and the business needs, such as energy savings, regulatory compliance, and service innovation.

Within the examined company, several A-BB and ML-BB were identified. For the selected A-BB, a suitable ML-BB was matched and reused to extend the analysis with ML–based anomaly detection. The Isolation Forest algorithm was used for this purpose and proved to be effective in detecting most dehumidification anomalies in the HVAC system. With the rule-based logic derived from the domain expert, the results of ML-detected anomalies were validated, which increased the trust in the model output, providing insights into inefficiencies within the HVAC system.

In the presented use case, the volume and complexity of data and anomalies were not very high, so the use of rule-based logic could be sufficient to detect the system's anomalous behaviour. However, over time, the data volume will increase; therefore, the rule-based logic could reach its limits and hinder the scalability and maintainability of the BB. Especially when reusing the ML-BB in the context of other use cases. To overcome this, we decided to go one step further and use ML algorithms that are able to handle high volumes of data and learn complex data patterns, so the efficiency, flexibility, and reusability of the ML-BB could be increased.

Finally, the modular structure of both A-BB and ML-BB made it possible to select the ML model in a targeted way based on the specific requirements of the use case. The selected ML-BB only needed minimal customisation, mainly adjusting thresholds to match the slightly different context. This shows that the BB can be reused across business services and use cases with minimal configuration effort.

Even though the results are promising, this study focused on a single dehumidification anomaly type, ignoring other anomaly types and use cases. As a result, the applicability and reusability to other use cases and contexts that are not closely related to those from which these BB are derived remain limited. The Isolation Forest model was only evaluated qualitatively, without standard performance metrics. Although the BB were designed for reuse, their application in other contexts has not yet been tested, and the matching between A-BB and ML-BB was done manually, which also does not guarantee that the selected ML-BB are the best fit.

Lastly, to enable the creation and selection of suitable BB that can be reused across different use cases, we envisioned in a separate paper currently under preparation that the context model should be available and analysed within the examined company, where the method for creating the A-BB and ML-BB is presented. The method already shows similarities to the TOGAF framework phases, but it does not explicitly describe them. Moreover, the context model covers all ArchiMate layers, and in the motivational layer, it provides the constraints and requirements of the HVAC system, which supports the selection of suitable BB for specific use cases. However, it does not define how the BB should be integrated into the larger portfolio. Since there are overlaps between the TOGAF phases and the method steps, it is possible to extend the method with additional steps similar to the TOGAF phases and notation elements of ArchiMate. These would guide stakeholders in setting business goals, selecting and implementing BB, analysing the cost and benefits of the implementation, and supporting them in migration planning. Furthermore, this will support the stakeholders responsible for implementation in identifying major projects and guiding them more effectively through the digital transformation. In this way, stakeholders are not only guided in the selection of BB but also in aligning them with business goals, planning their migration, and thus supporting digital transformation. Yet, this remains to be explored in future work.

In future work, we will focus on validating the results by reusing the BB in different contexts and use cases. Additionally, we plan to develop a repository to support the automated matching of BB. This will enable technical staff with limited IT expertise to select and apply appropriate BB for anomaly detection using ML applications.

Acknowledgments

The topics addressed in this paper are part of a joint research project between Dr. Diestel GmbH, the University of Rostock, the University of Applied Sciences Stralsund, and the University of Applied Sciences Wismar. The project is funded by the TBI Technologie-Beratungs-Institut GmbH (funding code: TBI-1-054-W-019).

Declaration on Generative AI

During the preparation of this work, the authors used Grammarly to check grammar and spelling. After using the tool, the authors reviewed and edited the content as needed to take full responsibility for the publication's content.

References

[1] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy and buildings 40 (2008) 394–398.
[2] N. Ivanovic, B. Nast, A. Reiz, K. Sandkuhl, Technologies for a diagnostic technique for HVAC systems using IoT and cloud-based architecture, in: 2023 International Interdisciplinary PhD Workshop (IIPhDW), IEEE, 2023, pp. 1–6.
[3] M. Bang, S. S. Engelsgaard, E. K. Alexandersen, M. R. Skydt, H. R. Shaker, M. Jradi, Novel real-time model-based fault detection method for automatic identification of abnormal energy performance in building ventilation units, Energy and Buildings 183 (2019) 238–251.
[4] W. Kim, S. Katipamula, A review of fault detection and diagnostics methods for building systems, Science and Technology for the Built Environment 24 (2018) 3–21.
[5] B. Nast, A. Reiz, N. Ivanovic, K. Sandkuhl, A modeling approach supporting digital twin engineering: Optimizing the energy consumption of air conditioning facilities, in: 2023 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C), IEEE, 2023, pp. 479–483.
[6] J. Luftman, Assessing business-IT alignment maturity, Communications of the AIS 4 (2000). doi:10.17705/1CAIS.00414.
[7] J. Luftman, K. Lyytinen, T. b. Zvi, Enhancing the measurement of information technology (IT) business alignment and its influence on company performance, Journal of Information Technology 32 (2017) 26–46.
[8] Y. E. Chan, B. H. Reich, It alignment: What have we learned?, Journal of Information technology 22 (2007) 297–315.
[9] F. Schlosser, H.-T. Wagner, T. Coltman, Reconsidering the dimensions of business-it alignment, in: 2012 45th Hawaii international conference on system sciences, IEEE, 2012, pp. 5053–5061.
[10] A. Yeow, C. Soh, R. Hansen, Aligning with new digital strategy: A dynamic capabilities approach, The Journal of Strategic Information Systems 27 (2018) 43–58.
[11] P. Gajardo, L. P. Ariel, The business-IT alignment in the digital age, in: The 13th Mediterranean Conference on Information Systems (ITAIS & MCIS), Naples, Italy, 2019.
[12] G. M. Jonathan, L. Rusu, W. Grembergen, Business-IT alignment and digital transformation: Setting a research agenda, in: Information Systems Development: Crossing Boundaries between Development and Operations (DevOps) in Information Systems (ISD2021 Proceedings), 2021.

[13] E. Piccinini, R. W. Gregory, L. M. Kolbe, Changes in the producer-consumer relationship-towards digital transformation, Wirtschaftsinformatik Proceedings (2015).
[14] M. Fitzgerald, N. Kruschwitz, D. Bonnet, M. Welch, Embracing digital technology: A new strategic imperative, MIT Sloan Management Review 55 (2014).
[15] C. Matt, T. Hess, A. Benlian, Digital transformation strategies, Business & information systems engineering 57 (2015) 339–343.
[16] G. Vial, Understanding digital transformation: A review and a research agenda, Managing Digital Transformation (2021) 13–66.
[17] D. R. Mckoy, R. C. Tesiero, Y. T. Acquaah, B. Gokaraju, Review of hvac systems history and future applications, Energies 16 (2023) 6109.
[18] P. W. Tien, S. Wei, J. Darkwa, C. Wood, J. K. Calautit, Machine learning and deep learning methods for enhancing building energy efficiency and indoor environmental quality–a review, Energy and AI 10 (2022) 100198.
[19] Y. Zhao, T. Li, X. Zhang, C. Zhang, Artificial intelligence-based fault detection and diagnosis methods for building energy systems: Advantages, challenges and the future, Renewable and Sustainable Energy Reviews 109 (2019) 85–101.
[20] C. G. Mattera, M. Jradi, M. R. Skydt, S. S. Engelsgaard, H. R. Shaker, Fault detection in ventilation units using dynamic energy performance models, Journal of Building Engineering 32 (2020) 101635.
[21] A. Arsanjani, L.-J. Zhang, M. Ellis, A. Allam, K. Channabasavaiah, S3: A service-oriented reference architecture, IT Professional 9 (2007) 10–17. doi:10.1109/MITP.2007.53.
[22] L. Thom, M. Reichert, C. Chiao, C. Iochpe, G. Hess, Inventing less, reusing more, and adding intelligence to business process modeling, in: International Conference on Database and Expert Systems Applications, 2008, pp. 837–850. doi:10.1007/978-3-540-85654-2_75.
[23] B. Molnár, P. Báldy, K. Menyhard-Balázs, Architectures of contemporary information systems and legal/regulatory environment, in: Proceedings of the 26th International Conference on Enterprise Information Systems, 2024, pp. 753–761. doi:10.5220/0012733600003690.
[24] Wikipedia contributors, Ipo model — wikipedia, the free encyclopedia, https://en.wikipedia.org/wiki/IPO_model, 2025. URL: https://en.wikipedia.org/wiki/IPO_model, accessed: 2025-07-10.
[25] Bundesrepublik Deutschland, § 71a GEG - Anforderungen an Gebäudeautomation, https://www.buzer.de/71a_GEG.htm, 2023. Accessed: 2025-07-16.