Multi-Partner Project: BIM-powered Environmental Data Agent for more Resilient and Trustworthy Data Centers

Oguzhan Herkiloğlu[1], Ali Serdar Atalay[1,4], İbrahim Arif[2], Salih Ergün[2,3], Alper Kanak[2,3]
[1]BİTNET Bilişim Hizmetleri, Ltd., Kastamonu, Türkiye; [2]ERGTECH SP. Z. O. O, Warsaw, Poland;
[3]Ergünler R&D Co.Ltd., Isparta, Türkiye, [4]AI4SEC OÖ, Tallinn, Estonia
Corresponding email: oguz@bitnet.com.tr

***Abstract*— This paper introduces an agent-based approach that semantically integrates the Building Information Model (BIM), Geographical Information System (GIS), and the Environmental Data Agent (EDA)-based optimization interface between Information and Operational Technology (IT/OT) for more trusted and resilient data centers. Using the cybersecurity-aware BIM-GIS-IoT data model facilitates the exchange of requirements and forecasts to optimize energy use, environmental impact, availability, and costs in data centers. At the core of this solution, the EDA securely mediates data exchange between IT and OT, translating IT resource consumption into energy metrics for effective optimization.**

Keywords*— *Environmental Data Agent, Energy Efficient Data Centers, BIM, Cybersecurity, IoT, Artificial Intelligence

I. Introduction

Data centers have been extensively established worldwide to satisfy the growing demand for Information and Communication Technologies (ICT) services, cloud computing, and IoT applications. According to a recent report, the 460TWh consumed by data centers in 2022 represented 2% of all global electricity usage [1]. To achieve net-zero emissions within the ICT ecosystem, existing data centers must undergo retrofitting to reduce the energy costs associated with cooling and ventilation systems [2]. A crucial element in achieving energy efficiency is the deployment of advanced sensor technology in data centers [3]. Sensors play a vital role in monitoring and managing the energy consumption of data centers, offering real-time insights into temperature, humidity, and airflow. By providing precise and continuous data, sensors enable dynamic adjustments to cooling systems, ensuring that energy is used efficiently and only where it is needed. Furthermore, sensors facilitate predictive maintenance by identifying potential issues before they lead to significant energy wastage or system failures. Sensors integrated with artificial intelligence and machine learning algorithms can optimize data center operations through advanced data analytics [4]. This paper addresses the increasing need for resilient, trustworthy, and energy-efficient data centers. The integration of the Building Information Model (BIM), Geographical Information System (GIS), and Environmental Data Agent (EDA) provides a comprehensive framework to optimize operations, reduce costs, and enhance cybersecurity. This work focuses on bridging the gap between Information Technology (IT) and Operational Technology (OT) domains, enabling dynamic adjustments to improve energy efficiency and operational trustworthiness. The challenges of managing sensor data, cybersecurity threats, and achieving net-zero emissions are explored in the context of real-world applications.

While sensors are crucial for optimizing the energy efficiency of data centers, the deployment of too many sensors can lead to several issues such as increased energy consumption, data overload and management complexity, higher costs, network bandwidth and latency, security and privacy concerns, calibration and accuracy issues, physical space constraints, integration challenges, and interference and redundancy problems [5][6]. As addressed in the Eco-Qube Project [7], there is a disconnect between Information Technology (IT) and physical infrastructure in data centers, which currently operate under the assumption of a static IT load. This static assumption is evident in Power Usage Effectiveness (PUE) calculations and virtualization practices that overlook physical infrastructure, leading to inefficiencies. IT servers are often underutilized, running at only 10-15% capacity, and many systems are over-provisioned, creating significant Green House Gas (GHG) emissions. Cooling systems must remain constantly available, as their only optimization signal is temperature, measured separately from IT operations.

From a cybersecurity perspective, integrating IoT devices into data centers can significantly enhance security through automated intrusion detection methods. These devices can continuously monitor the data center for physical and cyber threats, providing real-time alerts to authorities. On the other hand, this strategy transforms the data centers into more complex cyber-physical systems making them more vulnerable to cyber and physical attacks. As the number of internet-connected devices grows, handling the massive data streams requires powerful processing engines like complex event processing (CEP). Integrating machine learning (ML) and deep learning (DL) with CEP can further improve cybersecurity by enabling automated rule extraction and self-healing mechanisms [9]. From a physical resilience perspective, data centers are like huge digital living buildings where facility management, safety countermeasures against accidents or indoor/outdoor hazards, and protecting such critical infrastructures against malicious cyber or physical attacks are crucial. Recent studies utilize the Building Information Model (BIM) as a comprehensive data model that semantically interrelates the building components, IT infrastructure, data center, and the other operational external grids, i.e. electricity, water and gas pipelines, fiber lines, etc., [10][11]. BIM was used as the basis for modeling the complex nature of data center infrastructures as part of a digital twin architecture that enables effective and trusted monitoring of the facility and provides security and safety alerting and predictive protection against intended or unintended incidents [12].

This paper presents an optimization interface between IT and operational other technology (OT), as introduced in the Eco-Qube project, allowing for the exchange of requirements, demand profiles, and forecasts by relying on a cybersecurity and safety-aware BIM. This facilitates optimizations that balance energy use, environmental impact, availability, and costs, which are supposed to be achieved with less and more efficient use of sensors and well-calibrated services, addressing the efficiency gaps in current data center operations. The Environmental Data Agent (EDA) is the core concept of the proposed solution as it acts as an intermediary between IT and OT enabling secure data exchange without exposing OT to internet-related security risks. This system collects and processes data from both sides to optimize energy use, environmental impact, and costs. It translates IT resource consumption into electrical and thermal energy metrics, facilitating a common language for optimization.

The organization of the paper is as follows. Section II presents the EDA concept and the BIM-enabled semantic backend. Section III shows how this combined solution can be used in practical settings by giving examples from ongoing or finished research projects and demonstrations. Section IV concludes the paper and discussions on provisional studies.

II. BIM-Powered EDA Concept

The EDA acts as an intermediary between IT and OT, enabling secure and energy-efficient data exchange without exposing OT to internet-related security risks. This system collects and processes data from both sides to optimize energy use, environmental impact, and costs. It translates IT resource consumption into electrical and thermal energy metrics, facilitating a common language for optimization. BIM plays a complementary role in EDA as it provides a comprehensive building data model that is associated with OT. Originally, BIM does not address IoT and security aspects as it is designed for the design, planning, construction, and management of buildings. However, recent approaches extend BIM with GIS and IoT to address the living nature of buildings [13][14]. The following subsections present the EDA concept and the security- and safety-aware BIM-GIS-IoT integration.

A. *EDA for more energy-efficient data centers*

At the heart of the Environmental Data Agent (EDA) is a Prometheus server, which serves as the central repository for data collection and storage. Prometheus [15] is renowned for its robust time-series database and its ability to perform complex calculations, making it an ideal choice for monitoring and analyzing energy consumption patterns. The EDA leverages Prometheus's powerful querying capabilities to provide real-time insights and facilitate data-driven decision-making in energy management. By consolidating sensor data from various sources, EDA ensures comprehensive visibility and facilitates energy optimization.

This study enhances the traditional EDA by integrating it with Prometheus, enabling real-time monitoring and proactive decision-making. To support the diverse data collection needs, EDA integrates several auxiliary services tailored to specific data sources. For instance, the Intelligent Platform Management Interface (IPMI) is employed to gather critical hardware-level metrics such as power consumption and thermal readings directly from the IT/OT infrastructure. IT data like the operating system data, including CPU usage, memory utilization, and disk space consumption, is collected using standard system monitoring tools, ensuring comprehensive visibility into the system's performance. Additionally, the Modbus protocol is utilized to interface with cooling systems (OT), capturing data that is crucial for maintaining optimal operating conditions. Environmental and meteorological data (OT) are acquired through TCP connections, enabling the EDA to factor in external conditions that may influence energy usage.

The EDA acts as an intermediary between IT and OT systems, translating operational data into actionable metrics for optimization. The methodology employed to determine resilience improvements involves quantitative metrics such as downtime reduction and operational cost savings. By consolidating data from multiple sources into a unified platform, the EDA enables seamless integration and correlation of disparate data sets. This integrated approach allows for more accurate and granular analysis, facilitating proactive management of energy resources.

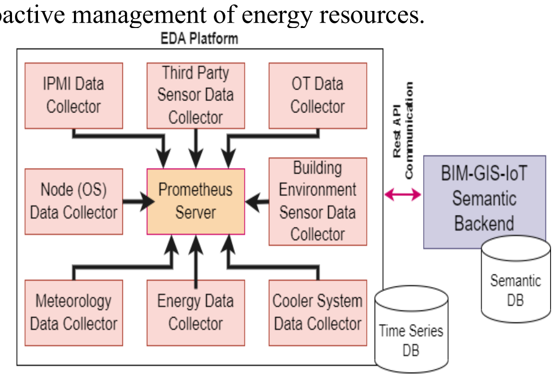

Fig. 1. The overall architecture of the EDA and BIM-GIS-IoT semantic backend

B. *Security- and safety-aware BIM-GIS-IoT integration*

The proposed data governance model is based on the semantic interlinking of four domain ontologies: I) BIM through IFC4.0.2.1 [16]; II) GIS through CityGML 3.0 [17]; III) IoT through Smart Sensor Networks (SSN) [18] and IV) cyber security derived from the ENISA threat taxonomy [19]. As seen in Fig. 2, BIM and GIS ontologies are linked with each other by the predicate “isSame” through their common concepts, IFCLocation and GMLLocation as well as IFCBuilding and GMLBuilding. Similarly, BIM and IoT are integrated over IFCSensor and IFCActuator (subclasses of IFCAsset) through the same predicate “isSame”. An event-based ontology is designed for the semantic model of cyber threats by relying on the ENISA threat taxonomy. This event-based ontology is linked with BIM and GIS (indirectly with IoT through IFC and CityGML) through the predicate “hasEvent”. A cyber incident is mapped with either a building asset via BIM (e.g., the cooling system in the data center building and its related IoT nodes) or an outdoor OT asset via GIS (e.g., ambient temperature sensor, surveillance camera, etc.). Such an ontology-based approach presents an efficient and modular data governance that can be used to store and make smart queries to analyze the multimodal IoT and services data. This semantic backend is implemented mainly by using Virtuoso, a semantic database and query engine, MQTT-based IoT communication and messaging backbone, Cesium, and Unity-based 3D visualization tools. To improve the resilience of the system against cyber-attacks, a Hardware Security Module (HSM), secure IoT gateways, and authentication services are implemented to enable encrypted IoT messaging and authorized access control to the EDA and its online services.

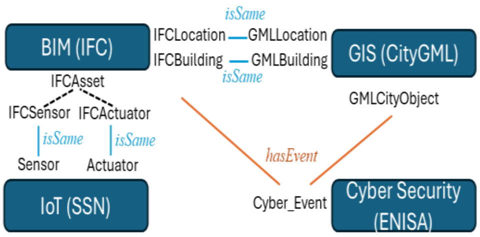

Fig. 2. Semantic integration of BIM, GIS, IoT, and cyber security

III. Practical Implications from Use Cases

This approach has been validated through the Eco-Qube[7] and BIMy [20] projects, demonstrating significant benefits in energy consumption and operational resilience. A retrofitted data center achieved a 15% reduction in energy usage while enhancing system availability. Detailed use cases highlight the seamless integration of BIM, GIS, and EDA, showcasing the potential for broader adoption. The system's direct impact on trustworthiness is demonstrated through improved monitoring, predictive maintenance, and enhanced cybersecurity measures.

A. Experimental Evaluation of EDA for Optimization of the Data Centers

The implementation of the EDA in optimizing energy-efficient data centers has yielded promising results. In a retrofitted air-cooled data center environment, similar to the study published in [2], the EDA system was thoroughly evaluated. This study underscores the challenges and opportunities in enhancing energy efficiency through retrofitting. The EDA's real-time data collection from various sensors and systems, including cooling infrastructure, server hardware, and operating system metrics, allowed for meticulous monitoring and control. By leveraging the EDA, the data center could dynamically adjust cooling strategies and optimize operations, resulting in significant energy savings while maintaining optimal performance.

The experimental setup utilizes EDA to manage cooling systems via the Modbus exporter, server hardware metrics via the IPMI exporter, and operating system metrics via the Node exporter. Additional data was gathered from various sensors through HTTP/TCP. The results mirrored the findings in [2], demonstrating a substantial improvement in energy efficiency. Specifically, the EDA-enabled data center achieved a 15% reduction in energy consumption, aligning with the benefits observed in the retrofitting study. The synergy between retrofitting efforts and advanced data monitoring through EDA underscores the critical role of integrated approaches in achieving energy optimization in data centers.

TABLE I. Evaluation of EDA for Data Center Optimization

Parameter	*Baseline (Pre-EDA)*	*Post-EDA Implement*	*Improvement (%)*
Energy Consumption (kWh)	150,000	127,500	15
Average Server Temperature (°C)	28	25	10.7
Cooling System Efficiency	3.0	3.5	16.7
Downtime (hours/year)	12	8	33.3
Operational Cost ($/year)	500,000	425,000	15

B. A Secure BIM-GIS-IoT Practice

A typical BIM model of a generic data center is designed by using Autodesk's REVIT tool. As seen in Fig. 3, a typical data center is equipped with sensors (motion, light, smoke, fire sprinkler, temperature, humidity, etc.), devices (power distribution units, server racks, uninterruptible power supplies, cooling systems, etc.) and grids for electricity, air, water, and gas. All indoor assets are modeled by using IFC classes (i.e. IFCAsset) or SSN objects. Similar to the authors' previous work [21], CityGML assets are used as asset classes for outdoor systems, sensors, and infrastructure in the close neighborhood of the data center building.

An example smart query is given below that is used to monitor the indoor and outdoor temperature sensor values. One can use these data and apply AI algorithms to calculate the energy efficiency of the data center by considering the OT and IT performance, estimated due dates for the maintenance of the sensor, devices, or building infrastructure, and detect anomalies in the observed system data. The BIM model file size is about 45MB and the average SPARQL response time is calculated as <1 seconds. Average memory and CPU consumption are measured as 4.2GB and 3%, respectively.

```
select ?asset_id ?stype ?sensor_result_time ?sensor_value
{?gml int: id 'input_id"^^xsd:string. ?ifc int:isA ?gml. ?ifc
int:globalID ?ifc_id; int:hasSpace ?space. ?space int:hasAsset
?asset. ?asset int:globalID ?asset_id. ?asset int:hasSensor
?sensor. ?sensor int:predefinedType ?stype.
FILTER(?stype = int:IndoorAmbientTemperature || ?stype =
int:OutdoorAmbientTemperature). ?sensor int:isSameSensorAs
?ssn_sensor. ?ssn_sensor sosa:madeObservation
ssn_observation. ?ssn_observation sosa:hasResult
?sensor_result. ?sensor_result sosa:resultTime
?sensor_result_time qudt-1-1:numericValue ?sensor_value.
FILTER(xsd:dateTime(?sensor_result_time)>=xsd:dateTime("
input_t1") && xsd:dateTime(?sensor_result_time)<=
xsd:dateTime("input_t2")) .}
```

PRIGM Cyber Security Solution Family [22] is used for holistic and point-to-point security in the EDA-compliant IoT backend. This family includes PRIGM Authentication Token, PRIGM Midi-1, and PRIGM Network HSM. PRIGM Token, a crypto-supported authentication device, provides authorized access to the BIM- and GIS-enabled IoT network. PRIGM Midi-1 serves as a secure IoT gateway for data exchange between the IFC, CityGML, and SSN nodes. Both devices collaborate with PRIGM Network HSM for active verification of edge devices on demand.

C. *Sensor Data Challenges*

While sensors play a critical role in optimizing data center operations, inaccuracies or malfunctions can compromise performance. Issues such as calibration errors, hardware malfunctions, and data latency can lead to suboptimal decision-making and reduced trustworthiness of the system. This paper proposes several strategies to mitigate these risks, including redundancy in sensor deployment, regular calibration schedules, and the application of AI-based algorithms to detect and correct anomalies. These measures ensure the reliability and accuracy of sensor data, which is critical for effective energy optimization, predictive maintenance, and cybersecurity monitoring. Addressing these challenges enhances the overall trustworthiness of the system and ensures reliable data for decision-making.

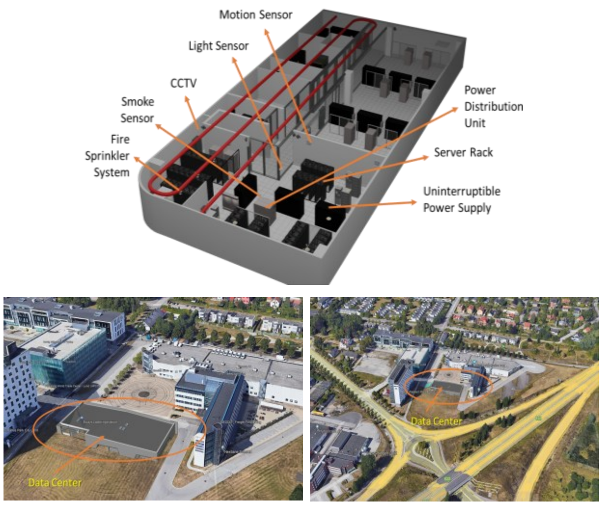

Fig. 3. Snapshots from the BIM-GIS-IoT visualization of a data center.

IV. Conclusion

This paper presents a combined approach that utilizes the Environmental Data Agent (EDA) concept for data center optimization by incorporating BIM-based building data. Multidimensional IT and OT data, associated with data center infrastructure and its connection with the outer world (e.g., GIS, IoT, energy grid, meteorological), are managed by IPMI, considering both energy efficiency and cyber resilience. The proposed approach demonstrates measurable improvements in energy efficiency, operational cost reductions, and enhanced resilience, showcasing its potential to address critical challenges in modern data centers. Future work will focus on extending this framework to more complex cyber-physical infrastructures and enhancing AI-powered tools for predictive analytics. Furthermore, efforts will be directed toward strengthening cybersecurity mechanisms and improving the integration of BIM-based semantic data governance tools to address the challenges of increasingly complex data center environments.

Acknowledgment

This work is supported in the Eco-Qube project (Horizon Europe Program, Grant ID: 956059), and relies on the BIMy project results (EUREKA-ITEA3, Grant ID: 16026).

References

[1] Electricity2024 Analysis and forecast to 2026, IEA Publications, International Energy Agency, https://iea.blob.core.windows.net/assets/6b2fd954-2017-408e-bf08-952fdd62118a/Electricity2024-Analysisandforecastto2026.pdf

[2] Kuzay, M., Dogan, A., Yilmaz, S., Herkiloglu, O., Atalay, A. S., Cemberci, A., ... & Demirel, E. (2022). Retrofitting of an air-cooled data center for energy efficiency. Case Studies in Thermal Engineering, 36, 102228.

[3] Mukherjee, D., Chakraborty, S., Sarkar, I., Ghosh, A., & Roy, S. (2020). A detailed study on data centre energy efficiency and efficient cooling techniques. International Journal, 9(5).

[4] Ahmed, K. M. U., Bollen, M. H., & Alvarez, M. (2021). A review of data centers energy consumption and reliability modeling. IEEE Access, 9, 152536-152563.

[5] Manganelli M, Soldati A, Martirano L, Ramakrishna S. Strategies for Improving the Sustainability of Data Centers via Energy Mix, Energy Conservation, and Circular Energy. Sustainability. 2021; 13(11):6114.

[6] Mondal, S., Faruk, F. B., Rajbongshi, D., Efaz, M. M. K., & Islam, M. M. (2023). GEECO: Green Data Centers for Energy Optimization and Carbon Footprint Reduction. Sustainability, 15(21), 15249.

[7] Eco-Qube Project, Artificial-Intelligence-Augmented Cooling System for Small Data Centres, https://ecoqube.org/.

[8] Ahmed, K. M. U., Bollen, M. H., & Alvarez, M. (2021). A review of data centers energy consumption and reliability modeling. IEEE Access, 9, 152536-152563.

[9] Alaghbari, K. A., Saad, M. H. M., Hussain, A., & Alam, M. R. (2022). Complex event processing for physical and cyber security in datacentres-recent progress, challenges and recommendations. Journal of Cloud Computing, 11(1), 65.

[10] Rybakova, A., & Kagan, P. (2020, June). Application of building information modeling in data center design. In IOP Conference Series: Materials Science and Engineering (Vol. 869, No. 2, p. 022006). IOP Publishing.

[11] Pogorelskiy, S., & Kocsis, I. (2023). Efficiency improvement with data center monitoring based on building information modeling on the facility management stage. Designs, 7(1), 3.

[12] Pinmas, P., Panyadee, P., Kaewmoracharoen, M., Khueankhan, S., & Wannachai, A. (2022, July). Towards Digital Twin Data Center Using Building Information Modeling and Real-Time Data Sensing. In 2022 37th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC) (pp. 908-911). IEEE.

[13] Kanak, A., Arif, İ., Kumaş, O., & Ergün, S. (2020, October). Extending BIM to urban semantic context for data-driven crisis preparedness. In 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 3813-3818). IEEE.

[14] Kanak, A., Arif, İ., Terzibaş, Ç., Demir, Ö. F., & Ergün, S. (2022, October). Bimyverse: Towards a semantic interpretation of buildings in the city and cities in the universe. In 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 1072-1077). IEEE.

[15] Open-source systems monitoring and alerting toolkit, built at SoundCloud, https://prometheus.io/.

[16] Industry Foundation Classes 4.0.2.1, V4.0 – Add.2 - Technical Corri.1,https://standards.buildingsmart.org/IFC/RELEASE/IFC4/ADD2_TC1/HTML/

[17] CityGML 3.0, https://github.com/opengeospatial/CityGML-3.0Encodings/releases.

[18] Semantic Sensor Network Ontology, OGC, W3, https://www.w3.org/TR/vocab-ssn/

[19] European Union Agency for Network and Information Security. (2016). ENISA Threat Landscape. ENISA. https://www.enisa.europa.eu/topics/threat-riskmanagement/threats-and-trends/enisa-threat-landscape/enisa-threat-landscape

[20] BIM in the City, BIMy project, https://itea4.org/project/bimy.html

[21] Kanak, A., Arif, İ., Terzibaş, Ç., Demir, Ö. F., & Ergün, S. (2022, October). Bimyverse: Towards a semantic interpretation of buildings in the city and cities in the universe. In 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 1072-1077). IEEE.

[22] PRIGM cyber security solution family, https://www.prigm.com.tr .