Scalable Digital Twins for Indoor Air Quality Management: A Review of Current Systems and Future Directions

Mehdi Ashayeri[1]

[1] *Southern Illinois University, Carbondale, IL*

Abstract. Indoor air quality (IAQ) management is moving from episodic assessment toward continuous, context-aware operation in which measurements, building-system data, and validated models are interpreted together. Within this transition, digital twins (DTs) offer a candidate scalable framework for linking calibrated sensing, building information modeling (BIM), semantic metadata, physics-based and hybrid IAQ simulation, data fusion, analytics, and control-oriented decision support. This review synthesizes DT and IAQ literature to identify the data, modeling, interoperability, and governance conditions that appear necessary for reliable deployment beyond isolated building pilots. The evidence reviewed suggests that scalable IAQ DTs are likely to depend on sensor calibration and uncertainty reporting, explicit spatial and system metadata, robust quality assurance/quality control (QA/QC), interoperable BIM and semantic representations, fit-for-purpose airflow and contaminant-transport models, edge-cloud processing, and privacy- and cybersecurity-aware governance. Although recent work demonstrates progress in IoT-enabled IAQ monitoring, BIM-linked asset representation, cloud/edge architectures, and machine-learning-based forecasting, many implementations remain pilot-scale because they lack standardized semantics, transparent validation, and documented operational performance across heterogeneous buildings. The review therefore envisions a layered IAQ DT architecture and staged implementation roadmap that could connect building-level pilots to federated urban-scale deployment.

Keywords: Digital twins; Indoor air quality (IAQ); Scalable systems; Semantic interoperability; Edge-cloud architecture

1. INTRODUCTION

Indoor air quality (IAQ) has become a major operational and public-health concern because people spend most of their time in enclosed environments, where inadequate ventilation can affect health [1], healthy-building conditions can influence performance and productivity [2], office pollutants can affect comfort and health [3], and ventilation is linked to airborne transmission risks [4]. Advances in low-cost IAQ sensing have expanded the feasibility of continuous indoor environmental monitoring [5]. However, low-cost IAQ and air-quality sensors still require careful performance evaluation [6] and field calibration against reference methods [7] before their data can reliably support operational decisions. In practice, IAQ monitoring often remains difficult to scale because raw sensor streams need to be calibrated, contextualized, and linked to building-system metadata before they can support portable analytics or control applications. Smart-building metadata research identifies heterogeneous point naming, subsystem representation, and manual data curation as major barriers to cross-building reuse [8]. Consequently, facility operators often lack integrated tools for translating environmental measurements into actionable intelligence for ventilation control, contaminant mitigation, energy optimization, and long-term operational planning. These limitations become increasingly important as IoT-enabled buildings and cities pursue energy-efficient operation [9], as global building-sector policy emphasizes sustainability and resilience [10], and as renovation policy promotes healthier and more efficient buildings [11]. Therefore, scalable IAQ management is likely to require more than dense sensing networks; it calls for interoperable cyber-physical infrastructures capable of integrating data acquisition, semantic building representation, simulation, analytics, and operational decision support into continuously updated systems.

Within this context, digital twins (DTs) have emerged as a promising paradigm for next-generation IAQ management because DT theory emphasizes synchronized computational representations of physical systems [12]-[13]. Broad DT research also identifies enabling technologies and open challenges for linking physical assets with data-driven models [14], while DT and big-data comparisons situate DTs within broader data-intensive smart-system architectures [15]. For buildings, semantic construction DT research indicates how digital representations can be enriched with semantic information [16]; separately, CPS-oriented DT literature in production systems supports the broader cyber-physical framing used here [17]. Unlike conventional monitoring systems that primarily support data visualization and descriptive reporting, DTs in the built environment are increasingly discussed as operational representations that combine asset information and sensor data [16], support built-environment applications and decision workflows [18], and may enable smart facility-management functions [19]. Built-environment DT reviews identify facility management, performance monitoring, and operational decision support as common application areas [18]-[19]. DT-enabled building-operation research also identifies fault detection and diagnosis as a major use case [20].

These studies also emphasize persistent challenges around data integration, interoperability, model updating, and organizational adoption [18]-[20]. For IAQ-oriented building operation, the most defensible near-term DT use cases include sensor-based monitoring supported by IAQ sensing research [5], HVAC fault detection and diagnosis supported by building-operation DT reviews [20], facility-management analytics supported by smart-FM literature [19], and simulation-assisted evaluation of airflow or contaminant dispersion supported by CFD and indoor-dispersion studies [21]-[22]. Indoor humidity and air-quality research further supports the need to connect environmental conditions with health-relevant IAQ interpretation [23]. Recent advances in smart-building metadata schemas improve application portability [8], semantic web technologies support AEC information integration [24], semantic BIM supports interoperable digital building twins [25], DT interoperability frameworks guide cross-platform integration [26], IFC supports standardized BIM exchange [27], and edge computing supports distributed processing architectures [28]. However, although DT research in buildings has expanded rapidly, built-environment and facility-management reviews still describe a gap between conceptual or application-specific demonstrations and mature operational deployment [18]-[19]. Digital-twin interoperability guidance further suggests that real deployments require sustained data integration, model updating, semantic consistency, and lifecycle-oriented system management [26].

Scalability in this review is grounded in wider urban-DT literature showing that city-scale twins are emerging across multiple built-environment systems, although many remain domain-specific, prototype-based, or framework-oriented. Water and water-energy studies show how DTs can link metering, hydraulic or energy-system models, resilience assessment, and operational decision support [85], [91]. Water-sensitive urban design reviews frame UDTs as platforms for monitoring, modeling, governance, and performance evaluation across complex urban water systems [95]. In solid-waste and public-realm regeneration, UDTs combine geospatial data, stakeholder participation, simulation, dashboards, and sustainability metrics to support planning decisions [86], [92]. Urban metabolism and New Urban Agenda studies extend this logic by connecting spatial representation, resource flows, sustainability indicators, and policy alignment at city scale [87], [96]. Smart-city DT reviews, including the corrected publication record, emphasize multi-layer urban applications and persistent data-governance challenges [88], [93]. Construction-sector research identifies data uncertainty, legacy systems, interoperability, security, and organizational capacity as implementation barriers [90]. Knowledge-graph and DT-BIM-GIS healthcare studies show that semantic integration and equity-sensitive spatial analytics are central to cross-domain urban decision systems [89], [94]. Together, this literature supports treating scalable IAQ DTs as governed, interoperable infrastructures for linking calibrated sensing, spatial context, model validity, lifecycle updating, and decision support.

A central challenge underlying this limitation is scalability, which remains insufficiently defined and systematically addressed within the existing IAQ DT literature. For IAQ digital twins, scalability can be defined not only as adding sensors or computing resources, but also as preserving data quality, semantic consistency, interoperability, synchronization, cybersecurity, and lifecycle maintainability across buildings. This definition aligns with smart-building metadata research on portable application semantics [8], digital-twin interoperability guidance on system integration and lifecycle concerns [26], and cyber-physical-system frameworks that treat composition, trustworthiness, lifecycle, data, timing, and boundaries as core system concerns [29]. Scalable IAQ DT ecosystems would need to connect heterogeneous sensors, building automation points, asset models, and analytics services through machine-readable metadata. Brick demonstrates how a semantic schema can represent building sensors, subsystems, and relationships to improve portability of smart-building applications [8]. Semantic BIM research indicates that building twins benefit from explicit information models rather than isolated data streams [25]. Digital-twin interoperability guidance similarly emphasizes interoperable interfaces and shared system information [26]. Because low-cost IAQ sensors can be affected by environmental conditions and sensor specifications [5], sensor-performance limitations and evaluation protocols should be considered [6].

Scalable deployments should also document indoor-air sampling strategy [30], field calibration procedures [7], reference comparison [6]-[7], and transparent reporting of measurement uncertainty before sensor data are used for forecasting or control. At portfolio scale, IAQ DTs can be envisioned as distributed cyber-physical systems in which building-level twins exchange standardized metadata [8], calibrated sensor streams [6]-[7], and operational analytics through interoperable interfaces [26]. This framing treats federation as a system-composition and data-governance problem within a broader CPS architecture [29] rather than merely a matter of adding more devices or computational capacity. As building systems become more connected, cybersecurity should also be treated as a scalability requirement rather than an add-on. Smart-building security reviews identify increasing protocol complexity, connected building-control systems, and limited empirical evaluation as persistent challenges [31]. Interoperability guidance and CPS frameworks likewise emphasize trustworthiness and system-boundary management for connected systems [26], [29].

Existing review studies have provided valuable insights into IAQ sensing technologies [5], IoT-enabled energy-efficient building and city systems [9], BIM and semantic web integration in AEC

[24], DT applications within the built environment [18], and DT-enabled smart facility management [19]. However, most reviews examine these domains independently and primarily emphasize technological capabilities rather than deployment scalability. Related review streams frequently concentrate on low-cost pollutant sensing and IAQ monitoring [5], sensor performance evaluation [6], or IoT-enabled energy-efficient building systems [9], rather than integrating these topics with semantic interoperability, DT synchronization, and operational scalability. Similarly, DT-oriented reviews often discuss conceptual architectures and enabling technologies in broad terms [14], semantic construction DTs [16], built-environment DT applications [18], or smart facility-management DTs [19], while giving limited attention to IAQ-specific contaminant dynamics, airflow modeling requirements, uncertainty management, or heterogeneous deployment conditions. Prior literature also rarely examines BIM interoperability [24]-[25], smart-building semantic modeling [8], edge-cloud orchestration [28], cybersecurity governance [31], QA/QC pipelines for sensing [5]-[7], and CPS-level scalability concerns [29] as integrated prerequisites for operational DT ecosystems. As a result, the reviewed literature presents many relevant technological advancements, but it does not yet provide a unified systems-level framework for understanding how these components collectively influence scalable IAQ DT deployment. This fragmentation has contributed to a persistent gap between proof-of-concept demonstrations and operational implementation across large and diverse building portfolios.

This review addresses that gap by reframing IAQ digital twins not as isolated visualization platforms or monitoring applications, but as a potential cyber-physical infrastructure perspective for coordinating sensing, semantics, modeling, computation, governance, and operational workflows. The intended contribution of this work is to examine scalability as an analytical lens for reviewing IAQ DT systems rather than to experimentally prove a universal deployment model. Rather than surveying DT technologies descriptively, this review synthesizes technical conditions identified in built-environment DT reviews [18], facility-management DT studies [19], interoperability guidance for digital-twin systems [26], and CPS frameworks for trustworthy connected systems [29].

Specifically, the review brings together literature on sensor calibration and uncertainty management [5]-[7], interoperable BIM and semantic metadata structures [8], [24]-[25], airflow and contaminant transport simulation [21]-[22], edge-cloud processing architectures [28], cybersecurity governance for connected building systems [31], cross-platform DT interoperability [26], and distributed multi-building integration through CPS system-composition and trustworthiness principles [29]. In addition, this study identifies recurring deployment bottlenecks that appear relevant to scalable IAQ DT adoption, including inconsistent metadata standards [8], insufficient validation and uncertainty-management methodologies [5]-[7], [30], limited operational benchmarking and deployment maturity in built-environment DTs [18]-[19], cybersecurity risks in connected building systems [31], and fragmented interoperability across DT software ecosystems [26]. These bottlenecks may help explain why many IAQ DT implementations remain closer to pilot-scale experimentation than to mature portfolio-wide operation. By positioning scalability as a systems-level review construct rather than a proven performance property, this paper offers a structured framework for assessing the maturity and potential deployability of IAQ DT ecosystems.

Finally, this paper envisions a layered conceptual architecture and staged implementation roadmap for scalable IAQ DT deployment that could help connect experimental research with operational adoption. The proposed perspective emphasizes modularity and machine-readable metadata representation through smart-building schemas [8], semantic interoperability through AEC semantic web technologies [24], interoperable digital building representation through semantic BIM [25], cross-platform DT interoperability through formal interoperability guidance [26], edge-cloud coordination through edge-computing architectures [28], and governance-aware system design through CPS lifecycle and trustworthiness principles [29]. Furthermore, the review highlights the importance of standardized validation procedures for low-cost IAQ sensing [5]-[6], indoor-air sampling strategy [30], field calibration and reference comparison [7], lifecycle-oriented deployment methodologies [26], [29], transparent uncertainty reporting [5]-[7], and cross-platform interoperability for more sustainable DT operations across heterogeneous building portfolios [26]. Based on the reviewed literature, the paper suggests that future IAQ DT systems may need to move toward federated, standards-based, and operationally resilient infrastructures capable of integrating sensing, simulation, analytics, and building management functions across multiple spatial and organizational scales. By synthesizing current research through the lens of scalability, this review aims to provide a structured foundation for future IAQ DT development and to clarify potential technical and organizational pathways for scalable implementation. Ultimately, scalable IAQ digital twins are presented here not as a proven end state, but as an emerging infrastructure vision for intelligent buildings, healthier indoor environments, and data-driven urban sustainability strategies.

2. REVIEW METHOD

This work is structured as a theory-building and synthesis-oriented literature review using selected representative studies rather than as a formal validation experiment or exhaustive systematic review. Guidance from evidence-informed reviews, information-systems literature reviews, systematic software-engineering reviews, and snowballing methods informed the selection and screening process [32]-[35]. Searches focused on combinations of terms including digital twin, indoor air quality, IoT sensing, low-cost sensors, BIM, building performance simulation, CFD, data fusion, machine learning, cyber-physical systems, smart buildings, and smart cities. Studies were retained when they contributed directly to one or more review layers: IAQ measurement, data quality, sensing infrastructure, fusion and analytics, BIM or semantic asset representation, physics-based simulation, occupant behavior modeling, cybersecurity, edge/cloud architecture, or urban scalability. This approach was used to develop a grounded conceptual synthesis and roadmap, not to claim that any proposed architecture has already been proven across all building types or deployment contexts.

Each retained source was evaluated for relevance, citation function, and scale of evidence. Relevance meant that the source offered substantive support for the sentence in which it was cited, extending beyond keyword-level association. Citation function meant that the source was classified as evidence for a particular layer: health motivation, sensor feasibility, data fusion, DT theory, BIM-enabled asset context, IAQ simulation, privacy and security, or urban-scale digital infrastructure. Scale of evidence distinguished conceptual frameworks, laboratory studies, building-scale deployments, and city-scale architectures. This classification is important because an IAQ DT that is plausible in one room or one building does not automatically generalize to a campus or municipality without additional work on interoperability, computation, governance, lifecycle maintenance, and data stewardship. The result is a deliberately conservative synthesis intended to clarify evidence boundaries and future system-design priorities.

3. DATA FOUNDATIONS FOR IAQ DIGITAL TWINS

A. Sensing and Data Acquisition

IAQ DTs begin with measurement, but measurement quality depends on how sensors are selected, calibrated, positioned, maintained, and interpreted. IoT and wireless sensor network research provides the communication foundations for dense sensing, while IAQ-specific sensor studies show both the promise and limitations of low-cost IAQ sensors [36]-[40]. Typical IAQ variables include particulate matter, carbon dioxide, temperature, relative humidity, carbon monoxide, nitrogen oxides, ozone, and volatile organic compounds, but not all variables have the same accuracy, response time, or stability when measured by inexpensive sensors. A review-based IAQ DT framework therefore should encode sensor metadata, calibration history, uncertainty, sampling interval, location, and environmental conditions alongside the measured value [5]-[7], [30]. At urban or community scale, data-driven urban building energy and microclimate studies suggest that outdoor drivers, building-stock attributes, and contextual metadata may enhance the interpretation and analysis of building operational energy demand [41]-[43]. Without those metadata, a platform may display IAQ measurements but remain unable to distinguish a real IAQ event from sensor drift, placement error, blocked airflow, or communication failure.

Data acquisition architecture is best matched to the decision being supported. Real-time ventilation control may require low latency, stable communication, and local fail-safe rules, whereas long-term exposure assessment can tolerate slower sampling if measurement uncertainty is well characterized. Dense sensing may reveal spatial gradients and localized pollutant events, but it also increases maintenance burden, network traffic, and data-quality risk. Conversely, low-density sensing may be easier to manage but can miss room-level exposure differences. The most defensible review-based recommendation is to design sensing as an integrated measurement system: identify the IAQ questions, define acceptable uncertainty, select sensors and reference checks accordingly, document sampling strategy, and implement routine validation. ISO indoor-air sampling guidance supports attention to sampling objectives and strategy [30], while low-cost sensor evaluation and field-calibration studies support the need for reference comparison and transparent uncertainty reporting before data are used for forecasting or control [6]-[7].

B. Data Fusion, Quality Control, and Analytics

Because IAQ data streams are heterogeneous, DT platforms are likely to require fusion and quality-control methods that reconcile multiple sensors, time scales, building systems, and contextual variables. Multisensor fusion literature provides a useful foundation for combining noisy or redundant observations, resolving conflicts, and improving state estimation [44]. In IAQ applications, fusion may involve combining fixed sensors with mobile measurements, linking pollutant readings to occupancy and HVAC operation, or integrating outdoor air data with indoor measurements. Machine-learning methods may support anomaly detection, missing-data imputation, sensor-drift identification, and short-term forecasting, with their strongest evidentiary role emerging when they are paired with data governance, calibration evidence, and physical understanding [45]-[47]. Hybrid physics-based machine-learning approaches (e.g., [48]-[97]) illustrate how predictive analytics can integrate sensor measurements and building physical conditions into a shared interpretive framework. Using social-sensing for multi-scale energy models (e.g., [49]–[50]) further demonstrate how urban context and human-related factors may be incorporated into data-driven or hybrid analysis. However, the transfer of these approaches to IAQ

DT deployment should still be considered a research direction unless validated against IAQ outcomes.

Quality control can be conceptualized at several levels of an IAQ DT architecture. At the edge, devices can screen impossible values, flag communication losses, and trigger local alarms when thresholds are exceeded. At the platform level, the DT can compare sensors against neighboring devices, expected physical relationships, historical baselines, and building-operation data. At the analytics level, models can quantify uncertainty and identify when predictions are extrapolating beyond the training domain. This layered approach is especially relevant for portfolio deployment, where many sensors may operate under different maintenance schedules and environmental conditions. The goal is not to eliminate uncertainty, but to make uncertainty visible, interpretable, and actionable. A scalable IAQ DT should therefore report data confidence, model confidence, and operational context together, with pollutant values interpreted through confidence measures and documented QA/QC assumptions.

C. Governance, Cybersecurity, and Privacy

Large-scale IAQ DTs raise governance issues because indoor environmental data can reveal information about occupancy, schedules, building use, maintenance practices, and potentially vulnerable populations. IoT security and privacy literature emphasizes that connected devices increase the attack surface through sensing nodes, networks, cloud platforms, application interfaces, and data-sharing arrangements [51]-[53]. Smart-building security reviews similarly identify connected control systems, protocol complexity, and limited empirical security evaluation as persistent challenges [31]. These risks are not abstract in IAQ management: malicious data manipulation could hide hazardous conditions, trigger inappropriate ventilation responses, or expose sensitive building-operation information. Security is therefore best framed as a design condition for IAQ DTs, with authentication, encryption, access control, device lifecycle management, audit trails, and incident-response procedures considered from the beginning. Privacy also warrants design-time consideration, with data minimization, aggregation, anonymization, and purpose limitation applied before large-scale data integration occurs.

Edge-cloud architectures may reduce both privacy and latency risks by processing sensitive or time-critical data closer to the source while reserving cloud resources for aggregation, long-term analytics, and cross-building comparison [28], [54]-[55]. For example, an edge device could detect a pollutant spike and initiate a local ventilation response without transmitting raw high-frequency occupancy-related data to a central platform. The cloud layer could then receive summarized events, calibrated trends, and metadata needed for benchmarking or planning. This separation supports scalability in principle because not every data stream needs to be centralized at full resolution. However, edge processing also introduces responsibilities for software updates, model synchronization, provenance tracking, local device security, and version control. Context-aware IoT and cloud-based CPS architectures support this layered interpretation, but IAQ-specific implementations should document which computations occur locally, which occur centrally, and how data provenance is preserved as information moves through the system [55]-[56].

4. SYSTEM FOUNDATIONS FOR IAQ DIGITAL TWINS

A. BIM, Semantics, and Building Context

A DT is likely to require more than sensor readings; it needs structured context that explains what the readings mean. BIM provides a representation of spaces, systems, components, and relationships, making it valuable for IAQ applications where pollutant behavior depends on rooms, zones, air-handling units, filters, dampers, windows, materials, and occupancy patterns [57]-[58]. Without this contextual layer, an IAQ platform may detect a CO2 change while lacking the information needed to distinguish whether the affected zone is a classroom, office, corridor, return-air path, or densely occupied meeting room. BIM and semantic models can also support asset maintenance by linking IAQ events to equipment, inspection histories, and control points. However, the reviewed literature also indicates that design BIM models are rarely ready for operations without cleaning, simplification, updating, and alignment with live data identifiers [24]-[25], [57]-[58]. For an envisioning paper, BIM is therefore best described as an enabling context layer rather than as proof of operational IAQ DT maturity.

For scalable IAQ DTs, semantic consistency is likely to be as important as geometric detail. Different buildings may use different naming conventions, sensor labels, HVAC control schemas, and facility-management platforms, so multi-building integration needs common data models or mapping layers. Smart-building metadata schemas and semantic BIM research support the idea that machine-readable relationships among buildings, zones, systems, sensors, and IAQ indicators improve application portability and interoperability [8], [24]-[25]. This requirement explains why many impressive single-building IAQ visualization interfaces do not automatically become scalable DTs: they may lack standardized semantics, governance rules, and comparable quality indicators. A practical IAQ DT roadmap should therefore prioritize a minimum operational information model that includes spaces, ventilation zones, sensor identity, equipment relationships, metadata, data-quality fields, and update responsibilities before adding advanced visualization or optimization features.

B. Physics-Based IAQ Simulation

Physics-based simulation remains a major evidence base for IAQ DTs because IAQ is governed by airflow, ventilation, deposition, filtration, chemical transformation, source strength, and contaminant transport. Ventilation-performance reviews, coupled building-energy and CFD studies, personalized ventilation simulations, and multizone airflow tools show how physical models can explain patterns that purely statistical models may miss [59]-[63]. For DT applications, physics-based models may support scenario testing, source identification, ventilation strategy comparison, and evaluation of interventions before implementation. CFD can provide detailed local airflow and contaminant distribution, while multizone models such as CONTAM represent building-level airflow and contaminant movement with lower computational burden. The appropriate model depends on the decision: room-level design may require CFD, whereas operational screening may be better served by calibrated reduced-order or multizone models.

Real-time IAQ DTs should be discussed with careful attention to model fidelity, computational cost, calibration burden, and operational purpose. High-fidelity CFD can be computationally expensive, difficult to calibrate, and impractical for continuous portfolio operation, especially when many buildings need to be updated simultaneously. A more scalable strategy suggested by the literature is to use physics-based models selectively: detailed simulations can generate insight, train surrogate models, or evaluate high-risk spaces, while lighter models support routine monitoring and decision support. Hybrid approaches are promising because physical constraints may improve machine-learning generalization and data-driven surrogates may approximate computationally expensive simulations [48]. However, this remains a proposed pathway unless models are validated for the relevant building, pollutant, occupancy condition, and ventilation regime. A DT documentation framework should therefore identify how models are calibrated, the conditions they represent, the uncertainty they carry, and the circumstances under which predictions fall outside the reliable operating envelope.

C. Occupant Behavior and Human-in-the-Loop Operation

Occupants influence IAQ through presence, activity, window operation, cleaning, cooking, equipment use, maintenance reporting, and responses to discomfort. Occupant-behavior research shows that human actions can strongly affect building performance and that behavior is variable, contextual, and difficult to predict deterministically [64]-[65]. Occupant-centered ambient-intelligence and linguistic studies also suggest that occupant feedback, perception, and sentiment can serve as human-in-the-loop signals when privacy protections are maintained (e.g., [66]-[71].) For IAQ DTs, this means that occupancy is best represented as a dynamic driver within IAQ processes, with schedules treated as contextual approximations rather than deterministic truth. Collecting occupant-related data also raises privacy and acceptability concerns, so the level of detail should be aligned with the decision need. Many IAQ decisions may use aggregated occupancy counts or schedule categories, avoiding reliance on personally identifiable movement data.

Human-in-the-loop operation is important because IAQ management involves trade-offs among health protection, comfort, energy consumption, noise, cost, and operational constraints. An automated system might increase ventilation to reduce CO2, but the decision may also affect humidity, thermal comfort, filtration load, or energy demand. Facility managers, occupants, and public-health stakeholders need explanations that connect recommendations to evidence and uncertainty. A useful IAQ DT should therefore be envisioned as decision support rather than autonomous proof of optimal action: it can provide interpretable alerts, scenario comparisons, confidence information, and traceable reasoning for control-related decisions. In practice, the DT may support decision-making at different levels: occupants need understandable feedback, facility staff need diagnostic information, and urban decision-makers need aggregated trends. Feedback loops can also record whether recommended actions were accepted, modified, or rejected by operators.

5. COMPARATIVE SYNTHESIS

This synthesis is organized in a deliberate sequence: first, a visual overview of the reviewed IAQ DT approaches; second, a data-layer matrix; and third, a system-layer matrix. The distinction is important because scalable IAQ DT performance cannot be inferred from measurement reliability alone or from system architecture alone. The data layer determines whether a platform can collect, validate, fuse, and protect IAQ measurements at appropriate spatial and temporal resolutions. The system layer determines whether those measurements can be interpreted through building context, simulation models, analytics, controls, and governance. Figure 1 provides a qualitative synthesis across sensing reliability, semantic integration, modeling capability, governance readiness, and scalability potential. Tables 1 and 2 translate that synthesis into requirements and limitations; they should be read as review-derived design considerations, not as evidence that any single approach has already achieved complete scalable IAQ DT deployment.

Figure 1. Qualitative synthesis of major IAQ DT approaches across key dimensions relevant to scalable deployment. The categories summarize broad patterns in the reviewed literature and indicate relative evidence and implementation readiness.

Table 1. Data-layer requirements for IAQ digital twins

Requirement	Main evidence base	IAQ DT implication	Scalability risk
Calibrated sensing	IoT, WSN, and low-cost IAQ sensor studies [36]-[40]	Measurements should be accompanied by sensor metadata, calibration history, location, sampling strategy, and uncertainty.	Uncalibrated devices can create misleading cross-building comparisons.
Data fusion and quality control	Multisensor fusion and ML literature [44]-[47]	Fusion may reconcile heterogeneous streams and flag missing data, drift, and anomalies when QA/QC rules are documented.	Large networks amplify drift, gaps, and metadata inconsistency.
Security and privacy	IoT security and privacy studies [51]-[53]	Authentication, encryption, access control, minimization, and audit trails should be embedded from the design stage.	Occupancy and operations data can expose sensitive building behavior.
Edge-cloud processing	Edge intelligence and cloud CPS architecture [54]-[55]	Local processing may reduce latency and privacy exposure while cloud services support aggregation and benchmarking.	Poor synchronization, versioning, or provenance tracking can fragment models and reduce trust in shared indicators.

Table 1 shows that sensing maturity is uneven: IoT and wireless sensor networks provide strong connectivity foundations, while IAQ-specific sensor studies define the calibration, uncertainty, and deployment constraints that determine whether measurements are decision-ready. Data-fusion methods are mature in principle, but they require careful adaptation to indoor environments where sensor placement, occupancy, and ventilation dynamics can produce localized IAQ patterns. Security and privacy are similarly mature in the broader IoT literature, yet they are often less fully embedded in IAQ DT implementation. The implication is that scalability should not be evaluated by data volume alone; it should also account for sampling strategy, calibration, metadata completeness, QA/QC, uncertainty reporting, lifecycle maintenance, and protected data workflows.

Table 2 shifts from data readiness to system capability. BIM and semantic DT research provide the contextual structure needed to connect IAQ measurements to spaces, assets, HVAC zones, and operational meaning, but they do not by themselves model contaminant transport or guarantee operational deployment. CFD and multizone simulation provide physical insight, but high-fidelity models are difficult to scale without reduction, calibration, or surrogate modeling. Machine learning may support forecasting and anomaly detection, but it can fail when training data are biased, poorly labeled, or weakly validated. Smart-city and cyber-physical-system architectures support federation across buildings only when local semantics, governance rules, and quality metrics are consistent enough to make comparisons meaningful.

Taken together, Figure 1 and Tables 1-2 suggest that the most defensible IAQ DT pathway is hybrid and layered: sensing provides observability, metadata and QA/QC establish trust, BIM and semantics provide operational context, physics-based models supply constraints, machine learning improves speed and pattern recognition, and governance protects comparability, privacy, and accountability. This integrated reading motivates the scaling discussion that follows, while preserving the review's central caution: these layers define a plausible implementation pathway rather than a universally proven solution.

Table 2. System-layer comparison for IAQ digital twins

System component	Supporting sources	Role in IAQ DT	Key limitation
BIM and semantics	BIM handbook and existing-building BIM review [57]-[58]	Can link sensor data to spaces, systems, assets, and operational meaning when operational identifiers are maintained.	Design BIM often lacks operational identifiers and live-data alignment.
Physics-based models	Ventilation, CFD, multizone, and CONTAM literature [59]-[63]	Can explain airflow and contaminant transport and support scenario testing within validated conditions.	High-fidelity models can be computationally costly and hard to calibrate.
Occupant behavior	Occupant-behavior modeling reviews [64]-[65]	Can represent occupancy and actions that influence sources and ventilation needs at an appropriate privacy-preserving resolution.	Fine-grained occupant data can create privacy and acceptability concerns.
Urban federation	Smart-city, real-time city, CPS, and context-aware IoT studies [55]-[56], [72]-[73]	May support cross-building comparison, governance, and city-scale services when standards and quality metrics are shared.	Heterogeneous systems require standards, quality metrics, lifecycle maintenance, and data stewardship.

6. SCALING FROM BUILDINGS TO LARGER SCALE IAQ DIGITAL TWINS

Scaling an IAQ DT from one building to many buildings involves qualitative changes in system architecture, governance, and evidentiary comparability. In this review, scalability refers primarily to repeatable deployment across building portfolios, campuses, districts, and other multi-building environments, while city-scale integration is treated as a possible extension when IAQ DT systems intersect with municipal platforms or public-health programs. Smart-city research remains useful because it addresses heterogeneity, institutional boundaries, data ownership, infrastructure variation, and governance complexity [72]-[73]. UBEM frameworks are also relevant because they link physical building stock and contextual indicators across groups of buildings and neighborhoods (e.g., [74]-[77].) A single building can often rely on manual setup, custom mappings, and local expert knowledge, whereas a portfolio-scale IAQ DT needs to integrate buildings with different ages, HVAC systems, sensor vendors, maintenance practices, occupancy patterns, and privacy requirements. This scale changes the technical problem: data volume increases, but the deeper challenge is comparability. Without common metadata and quality indicators, one building's pollution value may not be directly comparable with another's. Scalable IAQ DTs should therefore be understood as relying on standardization, federation, lifecycle maintenance, and transparent uncertainty reporting alongside computational capacity.

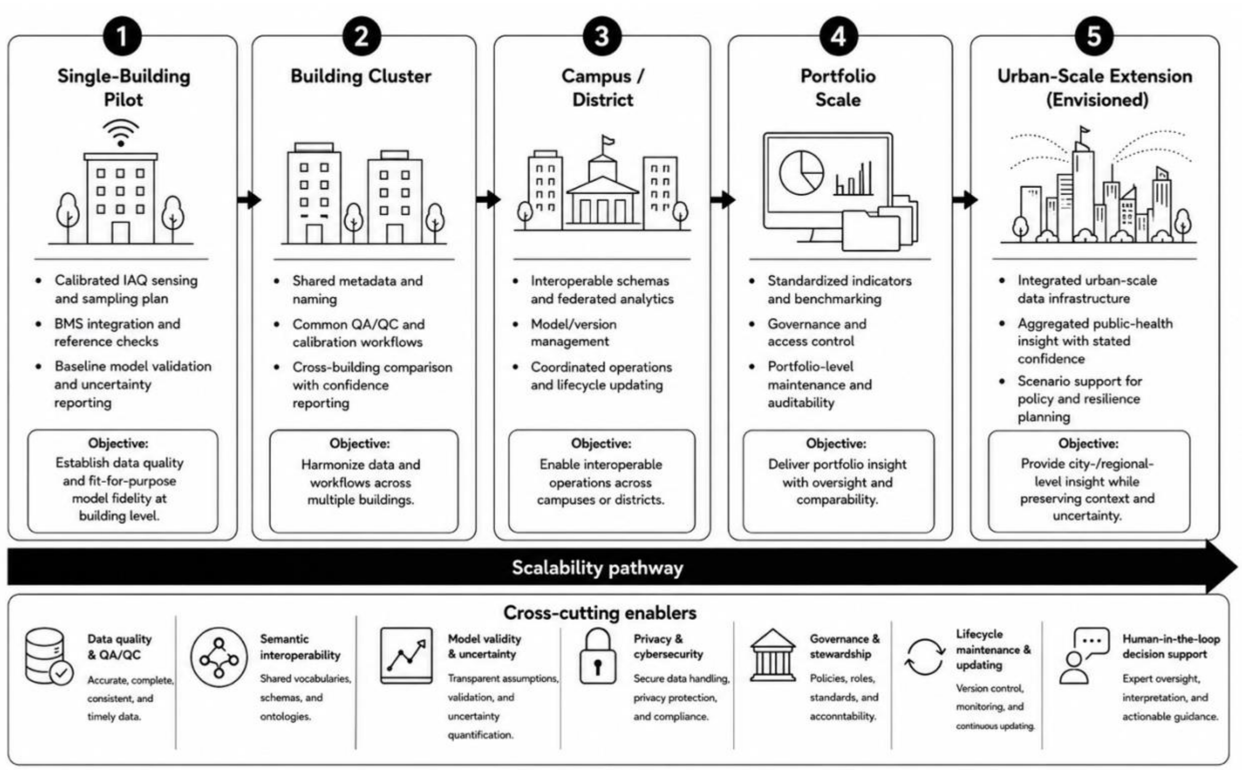

Figure 2. Scalability pathway for IAQ DTs, illustrating progression from isolated building pilots to deployment across building clusters, campuses, districts, portfolios, and urban-scale integration.

Reference architectures for cloud-based cyber-physical systems and context-aware IoT highlight the need for modular layers, contextual reasoning, and scalable data services [55]-[56]. For IAQ, these ideas translate into a possible federated portfolio architecture in which buildings retain local control over sensitive data and operations while sharing standardized indicators for benchmarking, public-health analysis, and policy support. Such a system could support applications such as identifying ventilation deficiencies across public buildings, prioritizing retrofits, assessing IAQ during wildfire smoke events, or coordinating healthy-building interventions. At higher levels of aggregation, the architecture needs to avoid false precision. An aggregate IAQ visualization interface should disclose differences in sensor quality, calibration age, sampling strategy, building type, or data completeness. The most credible portfolio-scale IAQ DT would present confidence-weighted indicators and allow users to trace aggregated patterns back to documented data sources and assumptions. Federation is therefore framed here as a promising governance and interoperability principle, not as evidence that a complete urban IAQ DT infrastructure has already been demonstrated.

This scalability pathway is synthesized in Figure 2, which conceptualizes progression from isolated building pilots to deployment across building clusters, campuses, and portfolios, while positioning urban-scale integration as a possible later extension. The figure should be interpreted as a staged roadmap for future development and evaluation rather than as a claim that each stage has already been implemented or validated in the reviewed literature.

7. RECOMMENDED ARCHITECTURE AND IMPLEMENTATION ROADMAP

Based on the reviewed literature, a scalable IAQ DT can be envisioned as a closed-loop, layered system with continuous feedback across sensing, modeling, decision support, and operation. The first layer is the physical building and IAQ source context, including spaces, occupants, HVAC systems, activities, materials, weather, schedules, and operational conditions. The second layer is sensing and acquisition, including calibrated IAQ sensors, building management system (BMS) data, occupancy indicators, airflow, energy use, outdoor air data, sampling metadata, calibration records, and uncertainty descriptors. The third layer is edge processing, where filtering, QA/QC, calibration checks, synchronization, local storage, event detection, privacy-preserving summarization, and software/model version control may occur before data are integrated into the wider platform [28], [54]-[55]. The fourth layer is the digital-twin core, where BIM or semantic models, asset relationships, topology, historical time-series data, and data-quality descriptors provide a structured representation of the building system [8], [24]-[27]. The fifth layer is modeling, analytics, and decision support, where physics-based simulation, machine-learning forecasts, anomaly detection, exposure assessment, risk scoring, scenario evaluation, IAQ visualization interfaces, alerts, and control recommendations are combined to support action. As shown in Figure 3, governance, privacy, cybersecurity, access control, auditability, and traceability are best treated as cross-cutting functions embedded across all layers of the architecture [26], [29], [31].

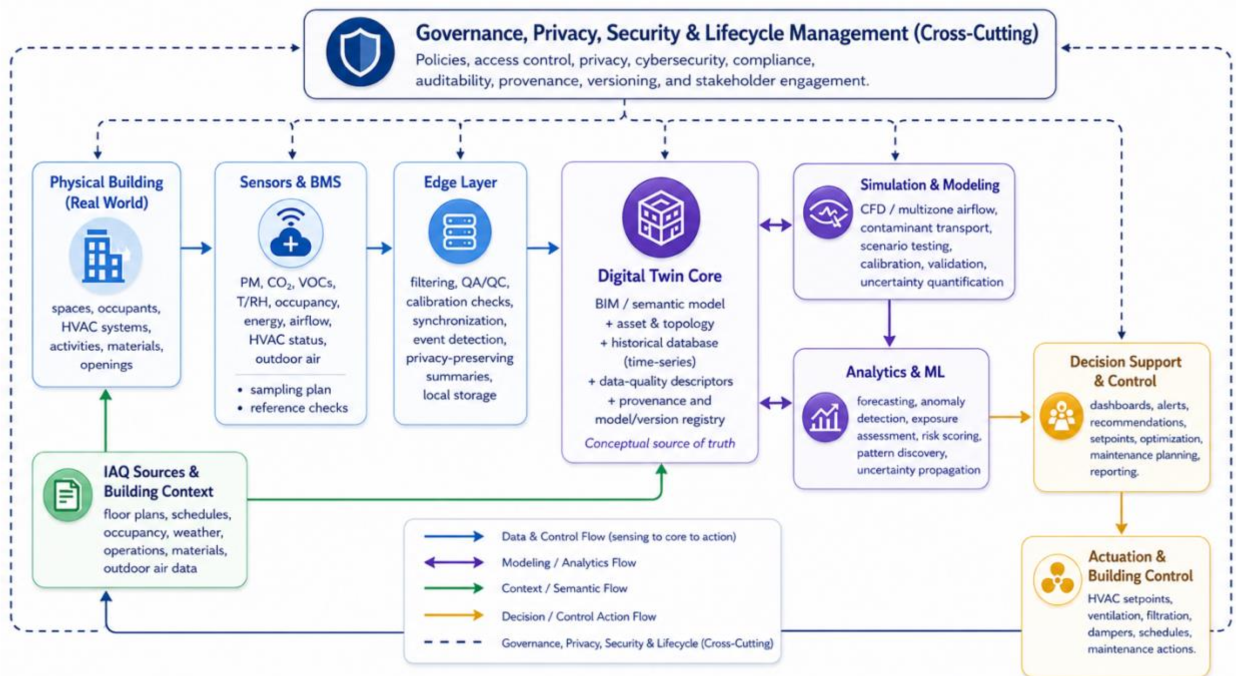

Figure 3. Recommended conceptual architecture for scalable IAQ DTs.

Implementation is best approached through staged maturity that precedes city-wide deployment. The first stage is a validated building pilot that focuses on calibrated IAQ sensing, BMS data integration, sampling strategy, metadata completeness, time synchronization, baseline IAQ performance, and user needs. The second stage integrates BIM or a simplified semantic building model so that sensor streams can be mapped to rooms, zones, HVAC systems, assets, schedules, operations, and data-quality descriptors. The third stage adds validated modeling, beginning with fit-for-purpose multizone, CFD, reduced-order, or hybrid models and expanding model fidelity only where uncertainty or operational risk justifies it. The fourth stage introduces predictive analytics and human-in-the-loop decision support, including forecasting, anomaly detection, exposure assessment, uncertainty visualization, alerts, and control recommendations. The fifth stage federates multiple buildings using common schemas, interoperability standards, governance rules, privacy and security protocols, and confidence-weighted indicators for benchmarking and policy support. Across all stages, lifecycle maintainability should be treated as part of scalability: sensors require recalibration or replacement, metadata mappings need updating, analytics models need versioning and revalidation, and governance rules may evolve as ownership, access, and reporting needs change [26], [29], [31].

Figure 4 translates this staged progression into a practical roadmap from building-level IAQ DT pilots to portfolio-scale and possible urban IAQ DT systems. It emphasizes that data quality, sampling strategy, semantic interoperability, model validity, decision usability, governance, privacy, security, stakeholder engagement, and lifecycle maintenance need to mature together as scale increases. The staged approach also makes evaluation easier because each maturity level can be assessed before additional complexity, automation, or broader deployment is introduced.

Figure 4. Staged implementation roadmap for moving from building-level IAQ DT pilots to portfolio-scale and multi-building IAQ DT systems.

8. DISCUSSION

The reviewed literature suggests that scalable IAQ DTs are best understood as coordinated cyber-physical and socio-technical infrastructures rather than as stand-alone monitoring dashboards. Figure 1, the data-layer synthesis, and the system-layer synthesis collectively indicate that sensing, metadata, modeling, governance, and operations need to mature together for IAQ DTs to be transferable beyond isolated pilots. Low-cost IAQ sensing can expand observability, but its value depends on calibration, sampling strategy, reference comparison, and uncertainty reporting [5]-[7], [30]. Similarly, semantic building metadata, BIM-based context, and DT interoperability guidance suggest that IAQ measurements become more useful when they are linked to spaces, assets, systems, and documented interfaces rather than treated as isolated time-series values [8], [24]-[27]. This interpretation is consistent with built-environment DT and facility-management reviews, which identify data integration, interoperability, organizational adoption, and operational decision support as recurring challenges rather than solved implementation problems [18]-[20].

The synthesis also points to several design trade-offs that should be made explicit in future IAQ DT work. One trade-off concerns model fidelity versus scalability: CFD and coupled airflow models can support detailed interpretation of local contaminant transport, while multizone, reduced-order, or hybrid models may be more feasible for routine operation across many buildings [21]-[22], [59]-[63]. A second trade-off concerns centralization versus federation. Cloud and edge architectures, context-aware IoT systems, and DT interoperability frameworks suggest that portfolio-scale systems may benefit from local processing, standardized interfaces, and federated indicators, but these choices introduce responsibilities for synchronization, provenance, version control, and lifecycle maintenance [26], [28]-[29], [54]-[56]. A third trade-off concerns observability versus privacy and cybersecurity. IAQ platforms may need occupancy, HVAC, and building-operation data to support interpretation, yet IoT and smart-building security literature shows that connected building systems can expand the attack surface and expose sensitive operational patterns if access control, privacy-by-design, and auditability are not embedded from the outset [31], [51]-[53]. Finally, automation should be balanced with human oversight because IAQ decisions involve comfort, health, energy, maintenance, and organizational constraints that cannot be reduced to a single technical objective [19], [64]-[67].

For implementation, the reviewed evidence suggests a staged and conservative pathway. Facility teams and platform developers are likely to gain more reliable value by first establishing a calibrated sensing and sampling plan, baseline IAQ indicators, QA/QC procedures, and BMS integration before adding advanced analytics or automated control [5]-[7], [30]. The next practical step is to map sensors and IAQ variables to rooms, ventilation zones, assets, equipment relationships, schedules, and data-quality descriptors through BIM or semantic models so that measurements can be interpreted across operational contexts [8], [24]-[25], [57]-[58]. Modeling and analytics can then be introduced in a fit-for-purpose manner: high-fidelity simulation may be reserved for critical spaces or scenario evaluation, while lighter models and ML methods may support anomaly detection, forecasting, and prioritization when their assumptions, training data, and uncertainty are documented [44]-[50], [59]-[63]. At multi-building scale, governance, access control, cybersecurity, maintenance responsibility, and confidence-weighted reporting should be treated as implementation requirements rather than optional additions [26], [29], [31], [51]-[54].

The limits of current evidence should therefore be stated carefully. The architecture and roadmap proposed in this review are synthesized from adjacent and partially overlapping literatures rather than validated as a universal IAQ DT solution. Sensor studies, semantic-modeling work, BIM literature, CFD and multizone modeling research, ML studies, edge-cloud architectures, and smart-city frameworks each support important parts of the proposed pathway, but few studies appear to demonstrate a fully integrated IAQ DT that is longitudinally validated across heterogeneous buildings, operational regimes, and governance settings. Built-environment DT reviews similarly report that many DT applications remain conceptual, pilot-scale, or domain-specific, with persistent challenges in interoperability, lifecycle updating, data quality, and organizational adoption [18]-[20]. This means that the figures and roadmap should be read as a review-based conceptual framework for structuring future deployment and evaluation, not as evidence that urban-scale IAQ DT performance has already been proven.

A focused research agenda can strengthen the field. First, future studies should report sensor specifications, sampling location and height, temporal resolution, calibration and reference-check procedures, data completeness, uncertainty, and maintenance history so that IAQ DT results can be compared across buildings [5]-[7], [30]. Second, model studies should document calibration datasets, validation periods, assumptions, failure modes, version history, and uncertainty propagation for physics-based, ML, and hybrid models [44]-[50], [59]-[63]. Third, common benchmarks are needed, including reference buildings, open or synthetic IAQ disturbance scenarios, standardized metadata templates, and performance metrics for both prediction and operational decision support. Fourth, longitudinal and multi-site studies should examine whether proposed architectures remain reliable across seasons, occupancy changes, HVAC modes, retrofits, and outdoor pollution episodes. Finally, governance research should examine data ownership, privacy expectations, public-health responsibilities, equity implications, operator trust, and human-in-the-loop decision-making when IAQ information is aggregated beyond a single building [31], [51]-[53], [56], [64]-[76]. These directions would help move IAQ DT research from promising conceptual integration toward evidence-informed, accountable, and scalable implementation.

A further direction for IAQ DT research is to extend the paper's scalability framework from technical deployment to context-aware environmental interpretation. This review has argued that scalable IAQ DTs require calibrated sensing, semantic building context, uncertainty reporting, validated models, lifecycle updating, and governance. At multi-building or urban scales, these requirements become more demanding because indoor conditions are shaped not only by building systems, but also by outdoor air, land-use context, occupancy behavior, and uneven social capacity to respond to environmental risk. Urban exposure studies therefore provide useful guidance by showing that air pollution concentrations and exposure estimates vary across spatial conditions, activity patterns, and population vulnerability [78]–[81]. For future IAQ DTs, this suggests that buildings should be represented as connected environmental nodes whose indoor conditions are influenced by external pollution, human use, and equity-relevant context. The same concern applies to the modeling layer: as IAQ DTs move from sensing to simulation, forecasting, and control-oriented recommendations, model outputs become useful only when their context of use is explicit.

Research on context-sensitive environmental modeling and AI-enabled performance-driven design provides a methodological precedent for documenting boundary conditions, assumptions, objectives, evaluation criteria, and trade-offs in computational decision support [82]–[83]. In IAQ DTs, this principle applies directly to airflow simulation, exposure estimation, predictive analytics, and ventilation recommendations, where uncertainty and decision logic should remain visible to operators and stakeholders. Finally, adaptive-urbanism theory supports the governance dimension of this framework by emphasizing flexibility under heterogeneous urban conditions [84]. This aligns with the staged roadmap proposed in this paper, where IAQ DTs progress from validated building pilots toward federated multi-building systems only when technical interoperability, evidence quality, lifecycle updating, and institutional responsibilities mature together.

9. CONCLUSION

This review suggests that IAQ digital twins may support more scalable indoor-air management when framed as governed cyber-physical infrastructures rather than as stand-alone monitoring dashboards. Using scalability as the organizing lens, the paper synthesizes how calibrated sensing, representative sampling, QA/QC, reference checks, uncertainty reporting, semantic building context, airflow and contaminant-transport modeling, analytics, and decision-support workflows need to be coordinated before IAQ DT outputs can be interpreted with confidence [5]-[8], [30]. In this framing, scalability is conditional rather than automatic: more sensors, higher model fidelity, or larger computing capacity are unlikely to be sufficient without interoperable metadata, data provenance, lifecycle updating, and explicit treatment of uncertainty.

The review further indicates that scalable IAQ DTs are likely to depend on interoperability, lifecycle management, governance, and human oversight as much as on advanced computation. BIM, semantic metadata, and DT interoperability guidance provide a basis for linking IAQ observations to rooms, ventilation zones, assets, systems, and operational identifiers, while edge-cloud and CPS-oriented architectures help frame distributed processing, synchronization, provenance, and version control across buildings [24]-[29], [54]-[56]. Physics-based, reduced-order, ML, and hybrid models should therefore be treated as fit-for-purpose components whose assumptions, calibration data, validation domains, and uncertainty are documented before they are used for control, benchmarking, or portfolio comparison [44]-[50], [59]-[63]. Governance, cybersecurity, privacy protection, and human-in-the-loop decision support remain essential because IAQ data can expose occupancy and operational patterns, and because IAQ interventions require balancing health, comfort, energy, maintenance, and organizational constraints [31], [51]-[53], [64]-[67].

Accordingly, this paper does not claim that broad IAQ DT deployment has been empirically proven; it offers a review-based conceptual pathway from building-level pilots toward federated multi-building and possible urban-scale applications. The proposed architecture and roadmap should be read as an evidence-informed framework for future implementation and evaluation, especially through longitudinal validation, transparent reporting, common benchmarks, stakeholder engagement, and governance mechanisms across heterogeneous buildings [18]-[20], [68]-[76]. Future work can test whether staged, layered, and governed IAQ DT approaches remain reliable across seasons, occupancy patterns, HVAC modes, retrofit conditions, and outdoor pollution episodes while preserving comparability, accountability, and trust.

REFERENCES

[1] Fisk, W. J. (2017). The ventilation problem in schools: Literature review. *Indoor Air, 27*(6), 1039–1051.

[2] Allen, J. G., & Macomber, J. D. (2020). *Healthy buildings: How indoor spaces drive performance and productivity*. Harvard University Press.

[3] Wolkoff, P. (2013). Indoor air pollutants in office environments: Assessment of comfort, health, and performance. *International Journal of Hygiene and Environmental Health, 216*(4), 371–394.

[4] Li, Y., et al. (2007). Role of ventilation in airborne transmission of infectious agents in the built environment: A multidisciplinary systematic review. *Indoor Air, 17*(1), 2–18.

[5] Sá, J. P., Alvim-Ferraz, M. C. M., Martins, F. G., & Sousa, S. I. V. (2022). Application of the low-cost sensing technology for indoor air quality monitoring: A review. *Environmental Technology & Innovation, 28*, Article 102551. https://doi.org/10.1016/j.eti.2022.102551

[6] Kang, Y., Aye, L., Ngo, T. D., & Zhou, J. (2022). Performance evaluation of low-cost air quality sensors: A review. *Science of the Total Environment, 818*, Article 151769.

[7] Spinelle, L., Gerboles, M., Villani, M. G., Aleixandre, M., & Bonavitacola, F. (2015). Field calibration of a cluster of low-cost available sensors for air quality monitoring: Part A: Ozone and nitrogen dioxide. *Sensors and Actuators B: Chemical, 215*, 249–257. https://doi.org/10.1016/j.snb.2015.03.031

[8] Balaji, B., et al. (2018). Brick: Metadata schema for portable smart building applications. *Applied Energy, 226*, 1273–1292.

[9] Al-Obaidi, K. M., Hossain, M., Alduais, N. A. M., Al-Duais, H. S., Omrany, H., & Ghaffarianhoseini, A. (2022). A review of using IoT for energy efficient buildings and cities: A built environment perspective. *Energies, 15*(16), Article 5991.

[10] United Nations Environment Programme. (2024). *2023 global status report for buildings and construction*. UNEP.

[11] European Commission. (2020). *A renovation wave for Europe: Greening our buildings, creating jobs, improving lives* (COM(2020) 662 final).

[12] Tao, F., Zhang, M., & Nee, A. Y. C. (2019). *Digital twin driven smart manufacturing*. Academic Press.

[13] Grieves, M., & Vickers, J. (2017). Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In F.-J. Kahlen, S. Flumerfelt, & A. Alves (Eds.), *Transdisciplinary perspectives on complex systems* (pp. 85–113). Springer.

[14] Fuller, A., Fan, Z., Day, C., & Barlow, C. (2020). Digital twin: Enabling technologies, challenges and open research. *IEEE Access, 8*, 108952–108971.

[15] Qi, Q., & Tao, F. (2018). Digital twin and big data towards smart manufacturing and Industry 4.0: 360 degree comparison. *IEEE Access, 6*, 3585–3593.

[16] Boje, F., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic construction digital twin: Directions for future research. *Automation in Construction, 114*, Article 103179.

[17] Negri, E., Fumagalli, L., & Macchi, M. (2017). A review of the roles of digital twin in CPS-based production systems. *Procedia Manufacturing, 11*, 939–948.

[18] Shahzad, M., Shafiq, M. T., Douglas, D., & Kassem, M. (2022). Digital twins in built environments: An investigation of the characteristics, applications, and challenges. *Buildings, 12*(2), Article 120.

[19] Hakimi, O., Liu, H., & Abudayyeh, O. (2024). Digital twin-enabled smart facility management: A bibliometric review. *Frontiers of Engineering Management, 11*(1), 32–49.

[20] Hodavand, F., Ramaji, I. J., & Sadeghi, N. (2023). Digital twin for fault detection and diagnosis of building operations: A systematic review. *Buildings, 13*(6), Article 1426.

[21] Marashian, S., Vadiee, A., Abouali, O., & Sadrizadeh, S. (2023). Enhancing indoor environmental simulations: A comprehensive review of CFD methods. In *Proceedings of the 64th International Conference of Scandinavian Simulation Society (SIMS 2023), Linköping Electronic Conference Proceedings, 200*, 365–371. https://doi.org/10.3384/ecp200046

[22] Dbouk, T., & Drikakis, D. (2022). Natural ventilation and aerosol particles dispersion indoors. *Energies, 15*(14), Article 5101.

[23] Wolkoff, P. (2018). Indoor air humidity, air quality, and health: An overview. *International Journal of Hygiene and Environmental Health, 221*(3), 376–390.

[24] Pauwels, P., Zhang, S., & Lee, Y.-C. (2017). Semantic web technologies in AEC industry: A literature overview. *Automation in Construction, 73*, 145–165.

[25] Roxin, A., Abdou, W., & Derigent, W. (2022). Interoperable digital building twins through communicating materials and semantic BIM. *SN Computer Science, 3*, Article 23.

[26] Digital Twin Consortium. (2021). *Digital twin system interoperability framework* (Version 1.0). Object Management Group.

[27] buildingSMART International. (2024). *Industry Foundation Classes (IFC) 4.3.2.0 standard* (ISO 16739-1).

[28] Perez, J., et al. (2022). Edge computing: A grounded theory study. *Computing, 104*, 2711–2747.

[29] Griffor, E. R., Greer, C., Wollman, D. A., & Burns, M. J. (2017). *Framework for cyber-physical systems: Volume 1, overview* (NIST Special Publication 1500-201). National Institute of Standards and Technology.

[30] International Organization for Standardization. (2004). *ISO 16000-1:2004, indoor air—Part 1: General aspects of sampling strategy*. ISO.

[31] Ciholas, P., Lennie, A., Sadigova, P., & Such, J. M. (2019). The security of smart buildings: A systematic literature review. *arXiv*. https://arxiv.org/abs/1901.05837

[32] Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

[33] Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *MIS Quarterly, 26*(2), xiii–xxiii.

[34] Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (Version 2.3, EBSE Technical Report EBSE-2007-01). Keele University; Durham University.

[35] Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering (EASE)* (Article 38, pp. 1–10). https://doi.org/10.1145/2601248.2601268

[36] Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks, 54*(15), 2787–2805. https://doi.org/10.1016/j.comnet.2010.05.010

[37] Gubbi, J., Buyya, R., Marusic, S., & Palaniswami, M. (2013). Internet of Things (IoT): A vision, architectural elements, and future directions. *Future Generation Computer Systems, 29*(7), 1645–1660. https://doi.org/10.1016/j.future.2013.01.010

[38] Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., & Cayirci, E. (2002). Wireless sensor networks: A survey. *Computer Networks, 38*(4), 393–422. https://doi.org/10.1016/S1389-1286(01)00302-4

[39] Yick, J., Mukherjee, B., & Ghosal, D. (2008). Wireless sensor network survey. *Computer Networks, 52*(12), 2292–2330. https://doi.org/10.1016/j.comnet.2008.04.002

[40] Spinelle, L., Gerboles, M., Kok, G., Persijn, S., & Sauerwald, T. (2017). Review of portable and low-cost sensors for the ambient air monitoring of benzene and other volatile organic compounds. *Sensors, 17*(7), Article 1520. https://doi.org/10.3390/s17071520

[41] Abbasabadi, N., & Ashayeri, M. (2024). Machine learning in urban building energy modeling. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.

[42] Worthy, A., Ashayeri, M., & Abbasabadi, N. (2025). Leveraging earth observational data products and machine learning to enhance urban building energy modeling (UBEM) with microclimate effects. *Sustainable Cities and Society, 130*, Article 106544. https://doi.org/10.1016/j.scs.2025.106544

[43] Worthy, A., Ashayeri, M., Julian, M., & Abbasabadi, N. (2025). Bridging the simulation-to-reality gap: A comprehensive review of microclimate integration in urban building energy modeling (UBEM). *Energy and Buildings*, Article 115392. https://doi.org/10.1016/j.enbuild.2025.115392

[44] Khaleghi, B., Khamis, A., Karray, F. O., & Razavi, S. N. (2013). Multisensor data fusion: A review of the state-of-the-art. *Information Fusion, 14*(1), 28–44. https://doi.org/10.1016/j.inffus.2011.08.001

[45] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[46] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444. https://doi.org/10.1038/nature14539

[47] Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science, 349*(6245), 255–260. https://doi.org/10.1126/science.aaa8415

[48] Ashayeri, M., & Abbasabadi, N. (2024). A hybrid physics-based machine learning approach for integrated energy and exposure modeling. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools*. Wiley.

[49] Abbasabadi, N., & Ashayeri, M. (2024). From tweets to energy trends (TwEn): An exploratory framework for machine learning-based forecasting of urban-scale energy behavior leveraging social media data. *Energy and Buildings, 317*, Article 114440. https://doi.org/10.1016/j.enbuild.2024.114440

[50] Abbasabadi, N., & Ashayeri, M. (2025). From tweets to energy trends (TwEn2): Social sensing-informed urban building energy modeling. *Frontiers in Energy Research*. https://doi.org/10.3389/fenrg.2025.1688348

[51] Roman, R., Najera, P., & Lopez, J. (2011). Securing the Internet of Things. *Computer, 44*(9), 51–58. https://doi.org/10.1109/MC.2011.291

[52] Sicari, S., Rizzardi, A., Grieco, L. A., & Coen-Porisini, A. (2015). Security, privacy and trust in Internet of Things: The road ahead. *Computer Networks, 76*, 146–164. https://doi.org/10.1016/j.comnet.2014.11.008

[53] Yang, Y., Wu, L., Yin, G., Li, L., & Zhao, H. (2017). A survey on security and privacy issues in Internet-of-Things. *IEEE Internet of Things Journal, 4*(5), 1250–1258. https://doi.org/10.1109/JIOT.2017.2694844

[54] Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., & Zhang, J. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE, 107*(8), 1738–1762. https://doi.org/10.1109/JPROC.2019.2918951

[55] Alam, K. M., & El Saddik, A. (2017). C2PS: A digital twin architecture reference model for the cloud-based cyber-physical systems. *IEEE Access, 5*, 2050–2062. https://doi.org/10.1109/ACCESS.2017.2657006

[56] Perera, C., Zaslavsky, A., Christen, P., & Georgakopoulos, D. (2014). Context aware computing for the Internet of Things: A survey. *IEEE Communications Surveys & Tutorials, 16*(1), 414–454. https://doi.org/10.1109/SURV.2013.042313.00197

[57] Sacks, R., Eastman, C., Lee, G., & Teicholz, P. (2018). *BIM handbook: A guide to building information modeling for owners, designers, engineers, contractors, and facility managers* (3rd ed.). Wiley.

[58] Volk, R., Stengel, J., & Schultmann, F. (2014). Building Information Modeling (BIM) for existing buildings: Literature review and future needs. *Automation in Construction, 38*, 109–127. https://doi.org/10.1016/j.autcon.2013.10.023

[59] Chen, Q. (2009). Ventilation performance prediction for buildings: A method overview and recent applications. *Building and Environment, 44*(4), 848–858. https://doi.org/10.1016/j.buildenv.2008.05.025

[60] Zhai, Z. J., & Chen, Q. Y. (2005). Performance of coupled building energy and CFD simulations. *Energy and Buildings, 37*(4), 333–344. https://doi.org/10.1016/j.enbuild.2004.07.001

[61] Gao, N. P., & Niu, J. L. (2005). CFD study on micro-environment around human body and personalized ventilation. *Building and Environment, 40*(6), 795–806. https://doi.org/10.1016/j.buildenv.2004.01.026

[62] Dols, W. S., & Polidoro, B. J. (2015). *CONTAM user guide and program documentation version 3.2* (NIST Technical Note 1887). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.TN.1887

[63] Wang, L., & Chen, Q. (2007). Theoretical and numerical studies of coupling multizone and CFD models for building air distribution simulations. *Indoor Air, 17*(5), 348–361. https://doi.org/10.1111/j.1600-0668.2007.00481.x

[64] Yan, D., et al. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings, 107*, 264–278. https://doi.org/10.1016/j.enbuild.2015.08.032

[65] Hong, T., Yan, D., D'Oca, S., & Chen, C. F. (2017). Ten questions concerning occupant behavior in buildings: The big picture. *Building and Environment, 114*, 518–530. https://doi.org/10.1016/j.buildenv.2016.12.006

[66] Abbasabadi, N., & Ashayeri, M. (2024). Occupant-driven urban building energy efficiency via ambient intelligence. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.

[67] Ashayeri, M. (2024). Decoding global indoor health perception on social media via NLP and transformer deep learning. In N. Abbasabadi & M. Ashayeri (Eds.), *Artificial intelligence in performance-driven design: Theories, methods, and tools* (1st ed.). Wiley.

[68] Ashayeri, M., Piri, S., & Abbasabadi, N. (2024). Exploring U.S. occupant perception toward indoor air quality via social media and NLP analysis. *Journal of Environmental Science and Public Health, 8*, 49–58. https://doi.org/10.26502/jesph.96120205

[69] Ashayeri, M., & Abbasabadi, N. (2024). Unraveling energy justice in NYC urban buildings through social media sentiment analysis and transformer deep learning. *Energy and Buildings, 306*, Article 113914. https://doi.org/10.1016/j.enbuild.2024.113914

[70] Ashayeri, M. (2025). An LLM framework for modeling U.S. thermal comfort perceptions. *engrXiv*. https://doi.org/10.31224/6913

[71] Ashayeri, M. (2025). Temporal and spatial dynamics of public sentiment toward indoor ventilation in New York City: Evidence from social sensing and large language models. *engrXiv*. https://doi.org/10.31224/6731

[72] Batty, M. (2018). Digital twins. *Environment and Planning B: Urban Analytics and City Science, 45*(5), 817–820. https://doi.org/10.1177/2399808318796416

[73] Kitchin, R. (2014). The real-time city? Big data and smart urbanism. *GeoJournal, 79*(1), 1–14. https://doi.org/10.1007/s10708-013-9516-8

[74] Abbasabadi, N., Ashayeri, M., Azari, R., Stephens, B., & Heidarinejad, M. (2019). An integrated data-driven framework for urban energy use modeling (UEUM). *Applied Energy, 253*, Article 113550. https://doi.org/10.1016/j.apenergy.2019.113550

[75] Abbasabadi, N., & Ashayeri, M. (2019). Urban energy use modeling methods and tools: A review and an outlook. *Building and Environment, 161*, Article 106270. https://doi.org/10.1016/j.buildenv.2019.106270

[76] Abbasabadi, N., & Ashayeri, M. (2021). Socioeconomic determinants of public health and residential building energy use in Chicago. In *2021 International Conference: 27th World Congress of Architects (UIA 2021 RIO), ACSA, Volume II*.

[77] Abbasabadi, N., Moroseos, T. F., Ashayeri, M., & Meek, C. (2026). AI-enhanced urban building energy modeling for health-driven decarbonization in vulnerable communities. *Preprints.org*. https://doi.org/10.20944/preprints202604.1576.v1

[78] Ashayeri, M., Worthy, A., & Abbasabadi, N. (2026). Predicting spatial inequities in particulate pollution exposure across Seattle areas using machine learning-enhanced land use regression. *SSRN*. https://doi.org/10.2139/ssrn.6764479

[79] Ashayeri, M., & Abbasabadi, N. (2022). A framework for integrated energy and exposure to ambient pollution (iEnEx) assessment toward low-carbon, healthy, and equitable cities. *Sustainable Cities and Society, 78*, Article 103647. https://doi.org/10.1016/j.scs.2021.103647

[80] Ashayeri, M., & Abbasabadi, N. (2021). Energy justice, indoor air quality, and community resilience against Covid-19 pandemic. In *Environments by Design: Health, Wellbeing and Place, 26.2* (pp. 1–6). AMPS. https://amps-research.com/wp-content/uploads/2022/08/Amps-Proceedings-Series-26.2.pdf

[81] Ashayeri, M., Abbasabadi, N., Heidarinejad, M., & Stephens, B. (2021). Predicting intraurban PM2.5 concentrations using enhanced machine learning approaches and incorporating human activity patterns. *Environmental Research, 196*, Article 110423. https://doi.org/10.1016/j.envres.2020.110423

[82] Luitjohan, S., Ashayeri, M., & Abbasabadi, N. (2022). An optimization framework and tool for context-sensitive solar-driven design using cellular automata (SDCA). In *2022 Annual Modeling and Simulation Conference (ANNSIM)* (pp. 593–604). IEEE. https://doi.org/10.23919/ANNSIM55834.2022.9859496

[83] Abbasabadi, N., & Ashayeri, M. (Eds.). (2024). *Artificial intelligence in performance-driven design: Theories, methods, and tools*. Wiley.

[84] Abbasabadi, N., & Ashayeri, M. (2017). Towards an adaptive urbanism beyond hard control: The theories of Johnson and Lefebvre. In *ARCC architectural research addressing societal challenges* (Vol. 1, pp. 257–262). CRC Press/Taylor & Francis Group.

[85] Pesantez, J. E., Alghamdi, F., Sabu, S., Mahinthakumar, G., & Berglund, E. Z. (2022). Using a digital twin to explore water infrastructure impacts during the COVID-19

pandemic. *Sustainable Cities and Society, 77*, Article 103520. https://doi.org/10.1016/j.scs.2021.103520

[86] Cárdenas-León, I., Koeva, M., Nourian, P., & Davey, C. (2024). Urban digital twin-based solution using geospatial information for solid waste management. *Sustainable Cities and Society, 115*, Article 105798. https://doi.org/10.1016/j.scs.2024.105798

[87] Geremicca, F., & Bilec, M. M. (2024). Searching for new urban metabolism techniques: A review towards future development for a city-scale urban metabolism digital twin. *Sustainable Cities and Society, 107*, Article 105445. https://doi.org/10.1016/j.scs.2024.105445

[88] Peldon, D., Banihashemi, S., Le Nguyen, K., & Derrible, S. (2024). Navigating urban complexity: The transformative role of digital twins in smart city development. *Sustainable Cities and Society, 111*, Article 105583. https://doi.org/10.1016/j.scs.2024.105583

[89] Tan, Y. R., et al. (2025). Beyond connected digital twins: Can digital twins really deliver sustainable cities? *Sustainable Cities and Society, 131*, Article 106596. https://doi.org/10.1016/j.scs.2025.106596

[90] Kineber, A. F., Singh, A. K., Fazeli, A., Mohandes, S. R., Cheung, C., Arashpour, M., Ejohwomu, O., & Zayed, T. (2023). Modelling the relationship between digital twins implementation barriers and sustainability pillars: Insights from building and construction sector. *Sustainable Cities and Society, 99*, Article 104930. https://doi.org/10.1016/j.scs.2023.104930

[91] Tavares, M., Pérez-Sánchez, M., Carravetta, A., Coronado-Hernández, O. E., López-Jiménez, P. A., & Ramos, H. M. (2024). Smart feasibility optimization of hybrid renewable water supply systems by digital twin technologies: A multicriteria approach applied to isolated cities. *Sustainable Cities and Society, 115*, Article 105834. https://doi.org/10.1016/j.scs.2024.105834

[92] Allan, M., Rajabifard, A., Chen, B., & Foliente, G. (2025). Utilising urban digital twins for sustainable urban regeneration: Melbourne's Greenline planning and assessment framework. *Sustainable Cities and Society, 131*, Article 106661. https://doi.org/10.1016/j.scs.2025.106661

[93] Peldon, D., Banihashemi, S., Le Nguyen, K., & Derrible, S. (2025). Corrigendum to "Navigating urban complexity: The transformative role of digital twins in smart city development" [Sustainable Cities and Society, 111, Article 105583]. *Sustainable Cities and Society, 132*, Article 106834. https://doi.org/10.1016/j.scs.2025.106834

[94] Dong, Y., Reffat, R. M., & Liang, R. (2026). Integrating digital twins, BIM, and GIS for urban health equity: Advancing smart and sustainable healthcare planning. *Sustainable Cities and Society, 144*, Article 107429. https://doi.org/10.1016/j.scs.2026.107429

[95] Altindas, G., Dane, G., & Baba, A. (2026). Urban digital twins and water sensitive urban design: A systematic literature review. *Sustainable Cities and Society, 145*, Article 107456. https://doi.org/10.1016/j.scs.2026.107456

[96] Diaz-Sarachaga, J. M. (2024). May urban digital twins spur the New Urban Agenda? The Spanish case study. *Sustainable Cities and Society, 114*, Article 105788. https://doi.org/10.1016/j.scs.2024.105788

[97] Ashayeri, M. (2020). *A hybrid data-driven simulation framework for integrated energy-air quality (iE-Aq) modeling at multiple urban scales* [Doctoral dissertation, Illinois Institute of Technology]. ProQuest Dissertations & Theses. No. 27956499.