

A semantically grounded Digital Twin Architecture for occupant-centric indoor lighting optimisation using a multimodal framework

Sreelakshmi Kavuthimadathil¹, Satish BK¹, Yacine Rezgui²

1-Welsh School of Architecture, Cardiff University

2-School of Engineering, Cardiff University

Corresponding Author: Sreelakshmikm150@gmail.com

Abstract—Indoor lighting is a major contributor to occupant health issues in the built environment, as occupants spend 90% of their time indoors. Poor lighting conditions can cause inefficiency, discomfort, and health problems, significantly contributing to the costs of absenteeism in countries. Lighting significantly influences mood, circadian rhythms, alertness, and productivity, and poor lighting can cause problems such as eyestrain, headaches, and sleep loss. The current building management systems treat building lighting as a static, schedule-driven utility rather than a dynamic, occupant-responsive system. Digital twin technology can shift lighting control from rule-based to adaptive, intelligence-driven when the essential layers, including the semantic, simulation, and learning layers, are properly integrated. This paper proposes a bottom-up, layered approach to a semantically grounded lighting DT architecture that can optimise lighting and close the control loop through rich, multimodal human feedback. The architecture proceeds from the input data and simulation upward through a semantic layer to a surrogate model trained on outputs from lighting simulation engines and ultimately to an adaptive machine learning control layer that closes the loop with occupant feedback. The paper presents a thorough literature review of existing gaps in the effects of lighting on human physiology and cognition, cognitive and subjective assessment instruments, and digital and semantic integration for human-centric buildings. The paper later analyses simulation engine data structures and their semantic mappings, identifies ontological gaps in the photometric and human behavioural domains, and proposes ways to address them. Key occupant parameters, illuminance and CCT preference, glare, luminous contrast of the screen, demographic profile and task context are grounded as typed inputs and outputs within the semantic layer. The framework is positioned as the foundation for experimental validation in an office cubicle environment.

Keywords—Digital Twin, Human-centric Lighting, Multimodal feedback loop, Building Ontology, Semantic integration.

I. INTRODUCTION

Indoor lighting in buildings acts as a biological regulator, shaping the human circadian system, hormonal cycles, cognitive performance, and affective state through both visual and non-visual pathways [1]. The melanopsin-based response of intrinsically photosensitive retinal ganglion cells within the human eye (ipRGCs) to short-wavelength light establishes a biophysical mechanism by which environmental illumination conditions influence melatonin suppression, cortisol rhythms, and sleep-wake regulation [2]. This is especially important, as humans spend 90% of their time indoors [3]. The melanopic equivalent daylight illuminance was hence introduced by the International Commission on Illumination as the principal metric for non-visual effects of lighting. These advances have helped reframe Indoor lighting as a multidimensional problem of occupant physiology and psychology rather than just a photometric problem.

Experimental work on the psychology and physiological impact has shown the scope for this reframing. Studies on EEG during exposure under different lighting conditions, measuring cortical power spectral density, have shown that δ - and β -band activity in the parietal region vary significantly during both rest and cognitive task stages [4]. Similarly, other

lighting studies have found that lighting could inhibit low-frequency activity and activate high-frequency bands, resulting in increased alertness and improved cognitive performance [5]. Eye-tracking data and gaze patterns are potential sources for measuring visual discomfort, which subjective questionnaires alone cannot provide [6]. Despite this evidence base, current lighting control systems in most office environments in the UK and around the world are controlled by time-based schedules, binary occupancy sensors, occupancy-based schedules, or simple daylight-linked dimming [7]. Many of these control strategies often neglect the occupant's state of well-being, current task type, demographic characteristics, or subjective preferences [8]. Previous literature has highlighted that factors including occupant position, daylight availability, glazing type in offices, gaze direction, and seasonal variation require highly individualised light exposure conditions that cannot be addressed by existing uniform control strategies [9,10]. The existing gap between what biophysical science recommends, and the systems delivered by building management systems presents both a research challenge and a public health opportunity.

Digital twin (DT) has emerged as a promising technology that can bridge this gap. The DT in the built environment

continuously synchronises a virtual model with its physical counterpart through sensor streams, enabling scenario simulation, predictive control and real-time optimisation [11]. In order to build a DT for optimising lighting that advances the current practise, it should be capable of ingesting heterogeneous occupant response data, physiological behavioural, and subjective data, interpret it within a semantic context that encodes room geometry, luminaire configuration, occupant demographics and circadian state, and translate it into closed-loop control decisions and use them to drive adaptive control decisions.

Such systems will require solving three interconnected problems that existing literature have addressed in isolation: i) the semantic integration of heterogeneous data structures from lighting simulation engines, BIM, IoT sensors, and physiological measurement systems ii) The building of surrogate model that are computationally tractable that emulates physically based lighting simulation at the scale of the DT, and iii) the grounding of the machine learning inputs and outputs in the ontologically typed concepts that enable reasoning, interoperability, and explainability as no existing building or IoT ontology provides the vocabulary necessary the photometric distributions, melanopic spectral observations, or multimodal structured occupant response profiles.

The present paper tries to address all three aspects through a unified architectural framework. The main contributions are: i) a bottom-up five-layer DT architecture for occupant-centric lighting optimisation, ii) a systematic mapping of simulation lighting engine data structures to building ontology classes, with identification of current gaps; iii) a proposed ontology extension filling three identified gaps; and iv) a delineation of the ML grounding strategy linking surrogate model outputs to occupant feedback in semantic context. The experimental validation of this framework in a real office environment is ongoing. The findings and validation will be published in subsequent work.

II. LITERATURE REVIEW

Lighting effects on Human Physiology and cognition

Recent literature on the evidence for the impact of lighting conditions on the health and well-being of occupants, especially on non-visual effects, has proven complex. Light conditions influence mood, sleep, alertness, and cognitive performance through physiological responses, including hormonal secretion, pupillary dynamics, cortical activation, and modulation of the autonomic nervous system [12]. A recent experimental study across multiple illuminance conditions found that a melanopic EDI of 150 lx optimised comfort and satisfaction, while 250 lx improved cognitive performance, and that a melanopic EDI of 694 lx reduced comfort [13]. This shows a non-linear, context-dependent relationship between light exposure and human comfort and performance. Another mixed method experiment across a $2 \times 3 \times 2$ factorial design (illuminance $\times$ CCT $\times$ time-of-day) found significant interactive effects, with 5000 K optimal at 300 lux in the morning and 6500 K most effective at 700 lux, while afternoon light preferentially improved subjective mood states [14]. The study also highlights that physiological and subjective alertness measures were sensitive to lighting variations, but behavioural performance on timed tasks was not. This nonlinear pattern makes fixed set-point control in

modern office buildings inadequate and highlights the need for more adaptive, occupant-centric lighting systems. Though several building standards have prescribed 250 Melanopic EDI lux at the eye level for four hours per day in regularly occupied spaces, this does not consider the task type and the physiological and psychological requirements needed for each task, nor does it consider the substantial inter-individual variability in visual comfort, circadian response, alertness, and subjective lighting preferences. Additionally, Field studies investigating seasonal and orientational-based variation in melanopic EDI have confirmed that gaze direction and room geometry can result in significant within-room variation in the circadian light exposure [15]. The availability of natural daylight further increases the complexity. A systematic review of daylight strategies in office spaces found that a single fixed strategy cannot meet illumination needs for various tasks at different times of day, especially in buildings with a compact plan and limited access to the outdoor environment. Dynamic control responsive to occupancy, sky conditions, and time of day is the only viable solution.

Occupant behaviour and lighting control

Reinhart (2002) proposed a predictive model of occupant interaction with lighting systems that uses annual profiles of occupancy and work-plane illuminance, combined with probabilistic switching patterns derived from field data, to simulate how different occupant types control electric lighting and blinds [16]. The model proposes 4 main user types and distinguishes between active and passive users. The model later became the foundation of energy prediction in several tools such as Daysim. However, the model is purely descriptive and only predicts how occupants behave with no information on their satisfaction, preference or subjective comfort [17]. This limitation resulted in moving towards other feedback-informed models. Field study research by Bellia et al. Looked at solar radiation, illuminance, and occupancy data to develop an adaptive lighting and blind control algorithm that iteratively learns from switching and blind adjustment events [18]. The study also reported that, while occupants consciously respond to glare and insufficient daylight, they frequently fail to return to the energy-efficient default once comfort is restored. An automated adaptive control can correct this pattern without imposing a burden on the user. Hence, there is a need for systematic research on multidisciplinary models integrating engineering, psychology, and computer science to build personal occupant models that can evolve using operational data [19].

Hence, there has been an increase in Machine learning approaches to occupant-informed lighting studies in recent years. Seo et al. have designed a supervised learning model that recommends illuminance levels and CCTs based on task type, fatigue level, and emotional state [20]. While other uses of ML applications in lighting design and control have shown the potential of GAN-based surrogate models, XGBoost-based glare control for kinetic facades, and neural network illuminance prediction models, demonstrating that data-driven approaches can also match the accuracy of physics-based simulation at a fraction of the computational cost.

Digital twins and semantic integration for buildings

The building Digital twin literature has highlighted that semantic interoperability is a precondition for operational intelligence. A multi-source data integration framework for building DTs has also shown that by combining the Building Topology Ontology (BOT) with Brick Schema, i.e., treating brick: Location as a specialisation of bot: Zone, resolves the semantic heterogeneity between geometric and systems representation, enabling a unified ontology that supports both spatial and sensor-actuator queries [21]. The integrated ontological framework, extended with SAREF for IoT device semantics and SSN/SOSA for sensor observation modelling, constitutes the present state of the art in building semantic integration [22]. Also, the Brick Schema has emerged as a scalable metadata standard for smart building applications, providing a structured, semantic framework for representing building assets, systems, and their operational relationships. It has an interoperability platform showing that BIM-to-Brick mappings can expose sensor and actuator data that IFC files alone cannot represent. SAREF4BLDG, an extension of SAREF, provides 72 classes and 179 object properties, including those for smart appliances, HVAC, and lighting devices [23,24].

The most SAREF-based energy modelling framework developed by Bjørnskov and Jradi is the operational example for DT, which is grounded in building ontologies. The model could generate an automated thermal model and real-time deviation tracking between physical and virtual systems using SAREF as the semantic backbone [25]. Though a few other frameworks address energy and thermal performance, no equivalent exists for lighting and occupant well-being. There is a gap in the literature in the semantic representation of photometric data, including the candela distribution, spectral power distributions, and melanopic EDI. These don't have any equivalence in any smart building ontology, a gap this paper identifies and begins to address.

Surrogate models and Reinforcement learning for building control

The existing physics-based simulations have computational limitations for real-time DT operation, though they are accurate. In contrast, learned approximations trained on simulation outputs have been effective at doing the same. Dey and Henze proposed an online reinforcement learning (RL) approach for building control in which data-driven surrogate models guide the RL agent during the early training phase, enabling faster, easier convergence than direct environment interaction [26]. A set of researchers has established surrogate models for lighting. Their study, based on a parametric design, presented the incidence operator method for determining optimal control-oriented surrogate models for illuminance on the work plane and daylight glare probability using weather data and adjustable material properties [27]. Researchers have also shown that the surrogate model serves as a DT simulation backbone rather than just a training aid. They have utilised energy plus, embedded in a feedback control loop that itself constitutes a DT, capable of predicting building response and optimising actuation over short time intervals using real-time sensor data [28]. The energyPlus Python energy management system (EMS) API, which exposes sensor handles and actuator handles for mid-simulation callback, enables precisely this pattern for electric lighting, daylight controls, and shade actuation [29].

III. PROPOSED ARCHITECTURE

Design principles

The proposed architecture for the digital twin is grounded in three design principles based on the literature review. First, a bottom-up data-first principle begins with available data and simulation capability, and the DT emerges as the terminal integration layer rather than the starting point. In contrast to the top-down DT approaches that impose a unified data model before the underlying data landscape is characterised, the bottom-up approach helps to reduce semantic mismatches and integration failures [11]. Second, a semantic grounding principle stating that every variable, including EEG feature vectors, eye-tracking fixation events, cognitive test response times, sensor readings, simulation outputs, Human factors, ML predictions, and control decisions that flow between layers, will carry an ontological type that enables reasoning and interoperability. Third, an occupant feedback closure principle will also be introduced. This control loop will not be closed by a fixed optimisation objective, but by real-time occupant satisfaction signals that continuously update the ML model's understanding of individual preferences and circadian needs.

Five-layer Architecture

The Architecture for the proposed Digital twin will consist of 5 layers, each with defined inputs, outputs, and interfaces (Figure 1)

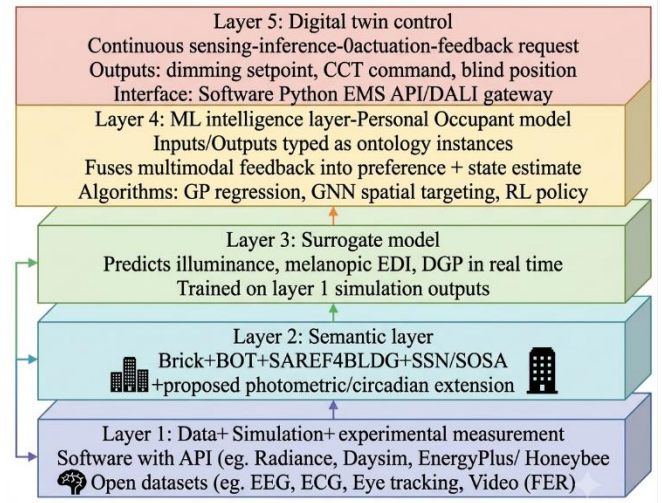


Figure 1: Proposed Five-layer Architecture

Layer 1-Data simulation, experimental measurements

Simulation sub-layer: The selected simulation software will provide physically validated ray-traced illuminance and luminance calculations [30]. The software can extend to annual climate-based simulation, generating a large number of illumination arrays, and annual metrics and extensions, such as the Lightswitch-2002 occupant behavioural model for daysim, can provide a probabilistic baseline for lighting and blind switching against which the personalised ML model is benchmarked. Software like EnergyPlus provides thermal and electrical lighting energy simulation via a Python EMS API that exposes mid-simulation sensor and actuator callbacks [29]. Honeybee/Ladybug Tools provides the HJSON format, a JSON-serialisable building model translating to

Radiance, EnergyPlus, and OpenStudio formats, serving as the geometric bridge layer [31].

Experiment measurement sublayer: The planned experimental phase will install a multimodal measurement protocol under controlled lighting conditions in an occupied office environment. This includes EEG measured using a 32-channel cap, Eye tracking using Tobii Pro, facial expression using the camera setup, cognitive tests including PVT, STROOP TASK, N BACK, and reading measured using the Psychopy test and subjective questionnaires like KSS, PANAS, and the Visual Comfort Questionnaire, administered at fixed intervals within each session. The participants are exposed to different lighting conditions, including a standard baseline, self-selected lighting conditions, and VTL biodynamic lighting delivered by a tunable luminaire with DALI protocol control, enabling precise, repeatable setpoint delivery and real-time monitoring.

Layer 2-Semantic Layer and HCL ontology Extension

The semantic layer assigns ontological types to all entities and data streams. Table 1 presents the complete simulation of ontology and experiment for ontology mapping.

Table 1: Data to Ontology Mapping

Data concept	Source	Data type	Ontology class
Room zone	BIM / HJSON	Name, area, volume	bot:Zone, brick:HVAC_Zone
Luminaire	IES/LDT file	Candela table, rated power	brick:Luminaire, saref:LightingDevice
Sensor point	Radiance / Daysim .pts	(x,y,z), normal vector	brick:Luminance_Sensor
Illuminance reading	Photosensor/simulation	Lux, timestamp	sosa:Observation brick:Illuminance_Sensor +
Lighting actuator	EnergyPlus / DALI	Dim 0→1, CCT setpoint (K)	brick:Lighting_System actuator
Occupancy	PIR / Wi-Fi	Boolean / count	brick:Occupancy_Sensor
EPW weather row	EnergyPlus	Dry-bulb, solar, hour	saref4envi:WeatherObservation
EEG observation	Amplifier	Band power ($\mu V^2/Hz$) per channel	hcl:EEGObservation
EEG band feature	Computed	$\delta, \theta, \alpha, \beta, \gamma$ power, ratio	hcl:EEGSpectralFeature
ECG observation	Wearable	R-R interval (ms), HRV indices	hcl:ECGObservation
Eye gaze event	Eye-tracker	(x,y) fixation duration, pupil d.	hcl:GazeObservation
Facial emotion	FER model	Probability vector (7 emotions)	hcl:FacialAffectObservation
PVT result	Software	Mean RT (ms), lapse count	hcl:CognitiveTestResult(PVT subclass)
Stroop result	Software	Interference score, accuracy	hcl:CognitiveTestResult (Stroop subclass)
KSS response	Survey	Integer 1–9, timestamp	hcl:SubjectiveRatingObservation (KSS subclass)
PANAS response	Survey	PA score, NA score	hcl:SubjectiveRatingObservation (PANAS subclass)
Visual comfort rating	Survey	Likert per dimension	hcl:SubjectiveRatingObservation (VC subclass)
Melanopic EDI	Computed / spectroradiometer	lux (α -opic) per CIE S 026	hcl:MelanopicObservation
Candela distribution	IES / LDT file	C-plane $\times \gamma$ table	hcl:PhotometricDistribution
Occupant profile	Demographic survey	Age, sex, chronotype, role	hcl:OccupantProfile

The proposed extension models measured and computed values as observation entities, while profile-related information is represented as individual instances. Each instance includes provenance metadata to record its generation process, the responsible source, and the timestep. This design ensures that every data point in the knowledge graph is fully traceable to its instrument, session, and occupant. This will enable audit, model refinement, and privacy-preserving aggregation.

Layer 3-Surrogate model

The surrogate model approximates full physical simulation at DT timescales. The inputs are a feature vector constructed from factors such as room geometry parameters, Luminaire configuration, and boundary conditions drawn from EPW. Outputs are predicted illuminance at each sensor point (Lux), Melanopic EDI at occupant eye position (Lux), and DGP (probability 0-1).

The training data is generated via systematic parameter sweeps using the selected software to assess photometric accuracy, covering the experimental room geometry across different luminaire configurations. The data volume target is $\geq 50,000$ simulation points. Three candidate architectures will be compared: 1) Gaussian process regression, XG Boost, and a shallow neural network. The selection criterion is the mean absolute error on a held-out test set of 10,000 points. The model uncertainty is taken to the DT control layer, i.e., when the predicted illuminance or melanopic EDI uncertainty exceeds a threshold, the DT requests a physical sensor reading to confirm before issuing a control command.

Layer 4-ML Intelligence: Personal occupant model

The ML intelligence layer learns a mapping from the current environmental state, as inferred from sensors and occupant context to predicted occupant response, enabling proactive rather than reactive control. The personal model for each occupant is a Gaussian process classifier trained incrementally on that individual's accumulated feedback history. The main inputs are a semantic feature vector that includes current illuminance, CCT, melanopic EDI, time of the day, day of the year, occupant chronotype score, current task type, and fatigue state estimated from KSS or EEG. The ML's output is the probability of satisfaction for each feasible lighting setpoint.

There are four main modalities that feed the model with different temporal granularities

- Continuous streams that run from layer 1 to layer 4 via layer 2. This includes the EEG β/α ratio, EEG HRV, eye tracking fixation entropy, and pupil diameter. These EEG-derived features and gaze-tracking observations will be semantically represented in the knowledge graph and accessed via SPARQL queries.
- Episodic measurements (experimental sessions): PVT lapse count, Stroop interference score, N back accuracy, and reading comprehension accuracy are semantically represented within the knowledge graph and are linked to the corresponding experimental session and associated lighting condition.
- Discrete self-reports: The subjective self-reported values of KSS, PANAS, and visual comfort are administered every 20-30 min during the sessions

and are stored and semantically structured in a knowledge graph together with their temporal and experimental context, enabling provenance-aware analysis.

- iv. Offline analysed signals (FER): Video frames are processed post-session using a ResNet-based FER model fine-tuned on AffectNet. The resulting 7-class emotion probability vectors are stored as semantically annotated observations, timestamped to enable synchronisation with physiological signals such as EEG, ECG, and Eye Tracking.

Feedback targeting: Active feedback requests (KSS and PANAS prompts) are directed by a Graph Neural Network-based spatial and temporal targeting algorithm, based on Abdelrahman and Miller's [32] model. The system identifies and selects the occupant, location, and moment most likely to improve model accuracy per unit of survey burden. This approach minimises fatigue from feedback requests while maximising the rate of personal model learning.

Multimodal optimisation:

The control objective is a weighted sum:

$$J(s) = w_1 E[M_{EDI}] + w_2 E[L_{wp}] - w_3 E[DGP] - w_4 E[P_{disc}] - w_5 E[E_{light}]$$

where $J(s)$ is the utility of a candidate lighting setpoint, M_{EDI} is the melanopic equivalent daylight illuminance, L_{wp} is the workplace illuminance, DGP is the daylight glare probability, P_{disc} is the predicted probability of visual discomfort, and E_{light} is the lighting energy consumption.

The weighting coefficients w_1 – w_5 reflect the relative importance assigned to circadian effectiveness, task illumination, glare mitigation, occupant comfort, and energy efficiency, respectively. These weights w_1 – w_5 are personalised per occupant and updated after each feedback batch. The optimisation is solved using Bayesian optimisation over the feasible setpoint space ($\dim \in [0.1, 1.0]$, CCT $\in [2700\text{ K}, 6500\text{ K}]$), with the personal model posterior as the acquisition function surrogate.

Layer 5- Digital twin control

The control layer in the digital layer operates in a closed-loop manner at 15-minute intervals. At each timestep, 1) the sensor readings, real-time environmental and occupant-related measurements are incorporated and added to the knowledge graph in the DT to update the current system. 2) The surrogate model is invoked with current boundary conditions to predict illuminance, melanopic EDI, and DGP. 3) The personal occupant model is queried for predicted satisfaction at each candidate setpoint. 4) The multi-objective optimiser selects the Pareto-optimal setpoint. 5) The command is issued via EnergyPlus Python EMS API (simulation mode) or DALI protocol (Physical deployment). Any differences between the predicted and measured illuminance trigger surrogate model recalibration; deviation between predicted satisfaction and measured KSS/PANAS triggers personal model update.

IV. DISCUSSION

The architecture proposed in the study advances the state of practice mainly across 4 domains: i) On data integration, it enables seamless integration of simulation outputs into a

knowledge graph by transforming native JSON-based building and simulation models into RDF, eliminating the need for any custom middleware and direct population, ii) it also proposes a semantic extension for representing photometric and circadian lighting data within the linked building data ecosystem, addressing a longstanding gap in the building ontology literature, iii) The framework also unifies physiological, behavioural, and subjective feedback streams within a common semantic framework, enabling multimodal learning, and finally, iv) the architecture uses ontological provenance via PROV-O enables differential privacy and data minimisation strategies. Sensitive data are processed locally, and only the derived semantic features are stored in the knowledge graph, with raw signals discarded after feature extraction.

At the time of writing, the experimental phase has just started; hence, the measurement protocol and data analysis strategies are intended methods rather than reported results. While the proposed framework establishes a unified, semantically grounded digital twin architecture, its primary limitation lies in its current stage of implementation, as the five-layer structure, surrogate models, and multi-objective optimisation weights (w_1 – w_5) are conceptually and heuristically defined rather than empirically validated. Furthermore, because the initial experimental protocol is designed around a single-occupant office cubicle setup, the framework's spatial generalizability, real-time performance against full Radiance simulations, and ability to handle multi-occupant control conflicts or survey fatigue in open-plan spaces remain to be verified in future empirical phases.

V. CONCLUSION

The present paper proposes a five-layer, semantically grounded Digital Twin architecture for occupant-centric indoor lighting optimisation. The notable contribution is the systematic integration of multimodal human feedback as first-class semantically typed data streams within the DT knowledge graph. Three main contributions are presented in the paper: i) mapping from simulation engine data structures and experimental measurement modalities to ontology classes, ii) a five-layer bottom-up architecture proceeding from one source simulation and experimental data, through a semantic layer and surrogate model, to a personal occupant ML model and closed-loop adaptive control. iii) an overview of experimental design specifying instrumentation, protocol and expected data structures for the upcoming empirical phase.

The proposed architecture places occupant neuroscience and cognitive performance within the semantic continuum of the built environment, integrating them into a single interoperable knowledge graph. Experimental validation later will address surrogate model training, personal model learning curves, and control effectiveness against circadian and cognitive performance targets, which will be published in later publications.

ACKNOWLEDGEMENTS

The authors acknowledge the European Union Horizon Europe research and innovation programme under the Marie 101202351 — DyLiOpt for financial support.



The authors also acknowledge the Spark behavioural lab for facilitating the data collection and access to facilities for experimental sessions.

REFERENCES

- [1] H. Xiao, H. Cai, and X. Li, "Non-visual effects of indoor light environment on humans: A review," *Physiology & Behavior*, vol. 228, p. 113195, 2021.
- [2] R. J. Lucas, S. N. Peirson, D. M. Berson, T. M. Brown, H. M. Cooper, C. A. Czeisler, M. G. Figueiro, P. D. Gamlin, S. W. Lockley, J. B. O'Hagan, L. L. A. Price, I. Provencio, D. J. Skene, and G. C. Brainard, "Measuring and using light in the melanopsin age," *Trends in Neurosciences*, vol. 37, no. 1, pp. 1–9, 2014.
- [3] N. E. Klepeis, W. C. Nelson, W. R. Ott, J. P. Robinson, A. M. Tsang, P. Switzer, J. V. Behar, S. C. Hern, and W. H. Engelmann, "The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants," *Journal of Exposure Analysis and Environmental Epidemiology*, vol. 11, no. 3, pp. 231–252, 2001.
- [4] L. Tong, N. Liu, S. Hu, M. Lu, Y. Zheng, and X. Ma, "Research on the preferred illuminance in office environments based on EEG," *Buildings*, vol. 13, no. 2, p. 467, 2023.
- [5] N. Zhang, C. Liu, C. Hou, Z. Wang, X. He, and W. Gao, "The application in electroencephalogram technology for investigating indoor lighting environments: A review," *Journal of Building Engineering*, vol. 98, 2024.
- [6] J. D. Blanco Cadena, T. Poli, M. Košir, G. Lobaccaro, A. G. Mainini, and A. Speroni, "Current trajectories and new challenges for visual comfort assessment in building design and operation: A critical review," *Applied Sciences*, vol. 12, no. 6, p. 3018, 2022.
- [7] J. Chen, J. An, D. Yan, and X. Zhou, "Future trends in intelligent lighting control systems: Integrated technologies, multi-system linkage and AI integration," *Building Simulation*, vol. 17, no. 11, pp. 1909–1932, 2024.
- [8] Z. Nagy, F. Y. Yong, M. Frei, and A. Schlueter, "Occupant centered lighting control: A user study on balancing comfort, acceptance, and energy consumption," *Energy and Buildings*, vol. 126, pp. 310–322, 2016.
- [9] M. G. Figueiro, B. Steverson, J. Heerwagen, K. Kampschroer, C. M. Hunter, K. Gonzales, B. Plitnick, and M. S. Rea, "The impact of daytime light exposures on sleep and mood in office workers," *Sleep Health*, vol. 3, no. 3, pp. 204–215, 2017.
- [10] N. G. Vásquez, R. F. Rupp, R. K. Andersen, and J. Toftum, "Occupants' responses to window views, daylighting and lighting in buildings: A critical review," *Building and Environment*, vol. 219, p. 109172, 2022.
- [11] W. Kritzinger, M. Karner, G. Traar, J. Henjes, and W. Sihm, "Digital twin in manufacturing: A categorical literature review and classification," *IFAC-PapersOnLine*, vol. 51, no. 11, pp. 1016–1022, 2018.
- [12] C. Liu, J. Li, X. Wang, Q. Zang, W. Wang, and W. Gao, "A review of subjective evaluation, physiological indicators and cognitive performance in indoor light environment: The role of illuminance and correlated colour temperature," *Journal of Building Engineering*, vol. 95, p. 110224, 2024.
- [13] H. Alhammadi and S. T. Amer, "The effects of lighting levels on perception, performance and physiological responses in offices," *Journal of Environmental Psychology*, vol. 107, p. 102801, 2025.
- [14] J. Sun, Z. Luo, Y. Gao, Q. Dong, X. Qi, et al., "The interactive effects of illuminance and correlated colour temperature on mood and alertness evaluated through psychological and physiological measures," *Building and Environment*, vol. 294, p. 114394, 2026.
- [15] M. Englezou and A. Michael, "Investigation of the daylight spectrum in an indoor environment using CIE S 026 melanopic metrics," *Lighting Research & Technology*, vol. 55, no. 8, pp. 690–711, 2023.
- [16] C. F. Reinhart, "Lightswitch-2002: a model for manual and automated control of electric lighting and blinds," *Solar Energy*, vol. 77, no. 1, pp. 15–28, 2004.
- [17] C. F. Reinhart and G. R. Newsham, "Impact of occupant behaviour on lighting energy use," in *Proc. Building Simulation 2009*, Glasgow, 2009, pp. 1143–1150.
- [18] L. Bellia, F. Fragiasso, and E. Stefanizzi, "Why are daylight-linked controls (DLCs) not so spread? A literature review," *Building and Environment*, vol. 106, pp. 74–85, 2016.
- [19] W. O'Brien, A. Wagner, M. Schweiker, A. Mahdavi, J. Day, M. B. Kjærgaard, S. Carlucci, B. Dong, F. Tahmasebi, D. Yan, T. Hong, H. B. Gunay, Z. Nagy, C. Miller, and C. Berger, "Introducing IEA EBC Annex 79: Key challenges and opportunities in the field of occupant-centric building design and operation," *Building and Environment*, vol. 178, p. 106738, 2020.
- [20] J. Seo, A. Choi, and M. Sung, "Recommendation of indoor luminous environment for occupants using big data analysis based on machine learning," *Building and Environment*, vol. 198, p. 107835, 2021.
- [21] X. Xie, N. Moretti, J. Merino, J. Y. Chang, P. Pauwels, and A. K. Parlikad, "Enabling building digital twin: Ontology-based information management framework for multi-source data integration," in *Proc. W78 2022*, 2022.
- [22] M. Elhajj, "Semantic foundations for digital twins: the contribution of ontological analysis," *Frontiers in Computer Science*, vol. 7, 2026.
- [23] R. Teymourzadeh and Y. Nakazawa, "A scalable and interoperable platform for transforming building information with Brick ontology," arXiv:2509.16259, 2025.

- [24] ETSI, TS 103 410-6 V1.1.1: SAREF4BLDG: Extension of SAREF for Buildings Domain. Sophia Antipolis: ETSI, 2019.
- [25] J. Bjørnskov, A. Badhwar, D. S. Singh, M. Sehgal, R. Åkesson, and M. Jradi, "Development and demonstration of a digital twin platform leveraging ontologies and data-driven simulation models," *Journal of Building Performance Simulation*, 2025.
- [26] S. Dey and G. P. Henze, "Reinforcement learning building control: an online approach with guided exploration using surrogate models," *ASME Journal of Engineering for Sustainable Buildings and Cities*, vol. 5, no. 1, p. 011005, 2024.
- [27] S. O. Weber, S. Subramaniam, and P. Leistner, "A parametric design integrated sampling and general training approach for optimal control oriented surrogate models of light-related quantities," *Journal of Building Performance Simulation*, 2025.
- [28] NREL, "EnergyPlus as a digital twin: real-time control and scenario simulation," National Renewable Energy Laboratory, Golden, CO, Tech. Rep., 2023.
- [29] US DOE, EnergyPlus Input-Output Reference and Python EMS Application Guide. Golden, CO: NREL, 2024.
- [30] G. Ward Larson and R. Shakespeare, *Rendering with Radiance: The Art and Science of Lighting Visualization*. San Francisco, CA: Morgan Kaufmann, 1998.
- [31] Ladybug Tools (M. Sadeghipour Roudsari, founder), "Honeybee Schema: HBJSON specification v1.x," GitHub, 2024.
- [32] M. M. Abdelrahman and C. Miller, "Targeting occupant feedback using digital twins: Adaptive spatial-temporal thermal preference sampling," *Building and Environment*, vol. 218, p. 109090, 2022.