buildings

MDPI

Systematic Review

Enhancing Systematic Review Efficiency with AIGC: Applications of Perception Data in Built Environment Audits

Anjun Tao [1], Zhijie Yang [1,2,*] and Wenbo Ou [1]

[1] School of Architecture, Southeast University, Nanjing 210096, China; taoaj@seu.edu.cn (A.T.)
[2] State Key Laboratory of Resources and Environment Information System, Institute of Geographical Sciences & Natural Resources Research, Chinese Academy of Sciences, Beijing 100871, China
* Correspondence: zhijieyang@ldy.edu.rs; Tel.: +86-19708878098

Abstract

With the growing use of human perception data streams in audits of the built environment, their value for enhancing objectivity and human-centeredness has become increasingly evident. This review synthesizes 63 publications through July 2024, providing a comprehensive analysis of perception data types, collection modalities and spatial strategies. This review introduces an Artificial Intelligence (AI)-enabled framework and utilizes Artificial Intelligence-Generated Content (AIGC) to assist literature retrieval and analysis, improving efficiency and transparency. The results indicate that heart rate and mood are currently the most frequently used perception data types in built-environment audits. Existing audit practices primarily focus on roads, green spaces, and residential areas at community and block-scale settings, with data choices varying by spatial typology. This review advances a systematic understanding of the application of perception data streams in built-environment audits and offers evidence-based recommendations for data collection, thereby providing stronger data support for future research.

Keywords: built environment audit; urban environment; AIGC; systematic review; generative artificial intelligence

check for updates

Academic Editor: James J.H. Liou

Received: 20 August 2025
Revised: 25 September 2025
Accepted: 5 October 2025
Published: 13 October 2025

Citation: Tao, A.; Yang, Z.; Ou, W. Enhancing Systematic Review Efficiency with AIGC: Applications of Perception Data in Built Environment Audits. *Buildings* **2025**, *15*, 3684. https://doi.org/10.3390/buildings15203684

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

With the rapid growth of urban populations and the intensification of climate and health risks, the built environment has become a critical part for improving population well-being and resilience. In developed economies such as the United States, the United Kingdom, Germany, and Japan, the mature urban systems consolidated during the early "golden age" [1] have been undermined by both external and internal shocks, including industrial hollowing-out, constrained municipal budgets, worsening public security and neighborhood decline [2,3]. In contrast, many developing countries face rapid population growth and incomplete urban governance and infrastructure systems, which have further accelerated urban decline and reinforced a self-perpetuating cycle of deterioration [4]. At the same time, Climate change is reshaping urban habitability. The built environment, which accommodates more than half of the world's population yet accounts for nearly 70% of global greenhouse gas emissions [5], is considered to be strongly associated with the mounting health risks among urban populations and has become a major global concern [6,7]. Therefore, the improvement, regulation, and planning of the built environment are crucial for enhancing human well-being. Research on the built environment spans multiple spatial scales, social groups, and extended temporal horizons, and the diversity of

built environments as well as the heterogeneity of research scales [8,9] complicate analysis and intervention. Conducting scientific assessments and selecting appropriate evaluation methods thus constitute the essential foundation. Built environment audits (BEA), as an important assessment tool, enable planners to systematically characterize existing conditions and identify priorities for planning and design improvements [7].

The built environment is defined as the man-made environment where people live, work, and play [10,11]. Encompassing both the physical structures and supporting infrastructure such as transportation, water, and energy networks. As the material, spatial, and cultural product of human labor and imagination [12], it was once primarily understood through its physical attributes. More recently, however, research has emphasized its strong impacts on quality of life and social equity [13], emotional cognition and mental health [14], and social order and crime prevention [15,16]. Therefore, BEA can also be conceptualized at two levels: a narrow and a broad sense. In the narrow sense, built environment audits focus on the objective measurement and evaluation of urban built spaces using standardized instruments—a practice traceable at least to the early 1980s. These audits quantify the quality and performance of the built environment against codified scales and national standards. For example, [8] conducted energy audits of buildings using Building Information Modelling, and [9] developed the Building Sync schema to standardize data audits of commercial buildings. Such approaches are predominantly technology-driven, aiming to produce precise, reproducible assessments of building and facility performance through standardized tools, models, and data protocols. In the broad sense, BEAs extend beyond physical metrics to integrate contextual, behavioral, and perceptual dimensions across multiple spatial scales. Rather than being confined to technical evaluations of discrete buildings or scenarios, it integrates multidimensional evidence, including psychological perceptions [14,15], physiological responses of plants and animals [17,18], and individual behavior [19–21]. Within this broad framework, scenario-driven practices that integrate intelligent methods are increasingly emerging. For example, some studies construct multiple scenarios and combine fuzzy expert systems (FESs) with self-organising maps (SOMs) to evaluate the appropriateness and trade-offs of intervention options: FESs formalise external expert knowledge and address uncertainty, while SOMs autonomously identify latent patterns and clusters in high-dimensional data. Taking cities such as Venice—characterised by a monofunctional focus on cultural tourism—as an example, related work compares three reuse scenarios for fortified heritage sites, integrating asset structural attributes, territorial context, and expert insights to propose actionable reuse pathways, thereby providing decision support in territories with functional imbalance and fragmented public policies [22]. Through interdisciplinary research designs, this approach yields comprehensive appraisals of the built environment [23,24].

In response to escalating climate risks, urban planning has moved toward demand-led allocation of nature-based solutions (NbS), where the siting and design of green-blue infrastructure are guided by mapped ecosystem service demand. By explicitly coupling demand with NbS supply, cities prioritize interventions in hotspots of heat exposure, pluvial flood risk, air-pollution burden, and deficits in access to restorative environments [25]. This trend underscores new requirements for built-environment audits—instruments must capture fine-grained environmental exposures, social vulnerability, mobility and use patterns, and human perception across scales to assess whether NbS are placed where they are most needed.

People's lived experiences constitute the fundamental criterion for judging environmental quality [26]. Given the diversity and complexity of built settings, traditional fixed evaluation scales struggle to deliver targeted assessments [27]. Incorporating human perception data into audits can therefore compensate for the limitations of purely objective

scales and offers distinctive advantages [28,29]. Although recent studies have begun to incorporate the perceptions of diverse experiencers into built environment audits [30,31], the elicitation and systematic use of perceptual information remain limited, and audit procedures tailored to different subject groups are underdeveloped. Moreover, existing work seldom analyzes perception data streams across time and lacks a coherent framework for deploying such data at multiple spatial scales of the built environment [28]. To address these gaps, we propose a systematic review to synthesize and generalize current research and to develop a multidimensional perception data framework spanning physiological, psychological, and behavioral domains.

AIGC refers to artificial intelligence technologies and applications capable of automatically generating digital content such as text, images, audio, and video [32]. Driven by advances in deep learning, natural language processing (NLP), computer vision, generative adversarial networks (GAN), and large language models (LLM) [33–36], AIGC has emerged as a significant trend in architecture and related fields [36,37]. In practice, AIGC supports generative design, qualitative evaluation, and a variety of applications such as integrating deep learning with physics-based urban energy modeling, automating layout generation with GANs, augmenting design ideation through AI image generation, and identifying construction risks via NLP [35,38–40]. These technological advancements highlight the potential of AIGC to transform systematic reviews by automating literature retrieval, enhancing data extraction, and expediting qualitative analyses, thereby increasing both efficiency and depth. However, current AIGC-generated content is still subject to distortion and inaccuracies [41,42], limiting its end-to-end application in literature synthesis. Meanwhile, the exponential growth of academic publications has made traditional manual review processes increasingly labor-intensive and time-consuming, often resulting in delayed dissemination of research [43–45], as well as inconsistencies caused by reviewer fatigue and subjectivity. Therefore, integrating AIGC with established research methodologies under appropriate oversight [46] offers a promising pathway to improve screening accuracy and synthesis efficiency, enhance the timeliness and quality of systematic reviews, and accelerate the diffusion of cutting-edge findings.

This study aims to integrate AI-generated content (AIGC) with manual review methods to efficiently screen and extract large volumes of research, systematically map the applications and evolution of human perception data in environmental audits, and evaluate the accuracy and feasibility of AIGC in supporting the review process. To this end, we develop an analytical workflow grounded in the SPIDER framework to accelerate qualitative synthesis using NLP and LLM. Specifically, the review will address:

1. The historical development of human perception data streams in the built environment.
2. Methods for deploying perception data streams and their spatial scales and applicability.
3. Future trajectories of built environment audits and their guidance for urban construction.

2. Application History of Perception Data Streams in Built Environment Audits

From Single to Multidimensional: The Application and Development of Perception Data in Built Environment Audits

The application of perception data in built-environment audits has evolved from manual observation and coding of behavior to the integration of multidimensional, streaming data. Early studies, such as those by Whyte [47], analyzed crowd behavior in urban public spaces using video recordings and manual coding, revealing the impact of environmental design on behavioral patterns. However, such research was largely limited to cross-sectional, static observation and could not capture individuals' physiological and psychological responses. With advances in sensing and computation, Di Rienzo, M. et al. [48]

monitored individuals' physiological data, such as heart rate and body temperature, using wearable devices to assess the impact of building environments on comfort. Clifton, KJ et al. [49] established a complete and reliable environmental audit method using the Pedestrian Environment Data Scan approach. Rainham, D. et al. [50] obtained a more comprehensive dynamic picture of the built environment affecting individual well-being through wearable GPS devices. These advancements marked a transition in the application of perception data from static behavioral analysis to dynamic physiological monitoring.

Subsequently, research gradually began to move towards integrated assessment using multidimensional perception data. Studies on built environment audits have shifted from single-source diagnostics to multi-objective, multi-perspective orchestration of information flows. Boarnet, MG et al. [51] incorporated various research projects on how density, street layout, mixed land use, pedestrian infrastructure, and a range of social and economic factors influence walking. By employing bidirectional and multivariate analytical methods to evaluate walkability in the built environment, they demonstrated that single-dimensional audits are insufficient for assessing complex environments. Meanwhile, Strath, SJ et al. [52] recruited older adults to collect geocoded objective environmental audit data and self-perceived environmental data, and integrated these with accelerometer-based physical activity behavior data for comprehensive assessment. The analysis of both subjective and objective environmental data further expanded the application of perception data in environmental audits. With big modern data analysis of behavioral patterns in populations, the importance of multi-source data integration in environmental audits has become increasingly apparent. The application of these perception data streams has collectively driven the development of built environment audits from a single dimension towards multidimensional and dynamic directions.

3. Materials and Methods

3.1. Approach to Searching and Selecting Literature

This review follows the principles of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [53], aiming to fill the gap in the selection of audit data extraction methods in built environment audits, thereby assisting future researchers in conducting more targeted audits of the built environment. Following the PRISMA 2020 protocol [54]: we conducted four stages: (1) identification and recording of relevant literature; (2) deduplication and title/abstract screening; (3) full-text eligibility assessment; and (4) inclusion with quality appraisal. Figure 1 presents the PRISMA 2020 flow diagram used in this study (the completed PRISMA 2020 checklist is provided in the Supplementary Materials).

In order to better tailor the application of AIGC technology and avoid blind selection when targeting qualitative research subjects, this study adopts the SPIDER framework as guidance prior to conducting the audit [55]. This review uses Web of Science and Scopus, the two largest English-language SCI databases, to retrieve relevant literature, with duplicates between the two databases removed. The publication date for retrieved studies is limited to between January 2000 and October 2024, as this window is sufficiently recent to align with contemporary data and methodological standards while long enough to capture robust trends and paradigm shifts since the early 2000s. The search strategy combines audit-related terms such as urban audit and virtual environment audit with human information data-related terms such as human behavior and mental health, searching within titles, abstracts, and keywords. The initial search string used in this study is as follows:

TypeScript = ((“built environment* audit*” OR “environment* audit” OR “urban audit” OR “virtual audit”) AND (“perception” OR “health” OR “behavior” OR “human” OR “physiology” OR “pedestrian”)).

Studies are excluded if the full text is not accessible, if they are not qualitative or quantitative research, if they have not been formally published, if they do not consider factors related to humans and built environment audits, if they are not published in English, or if they were published before the year 2000.

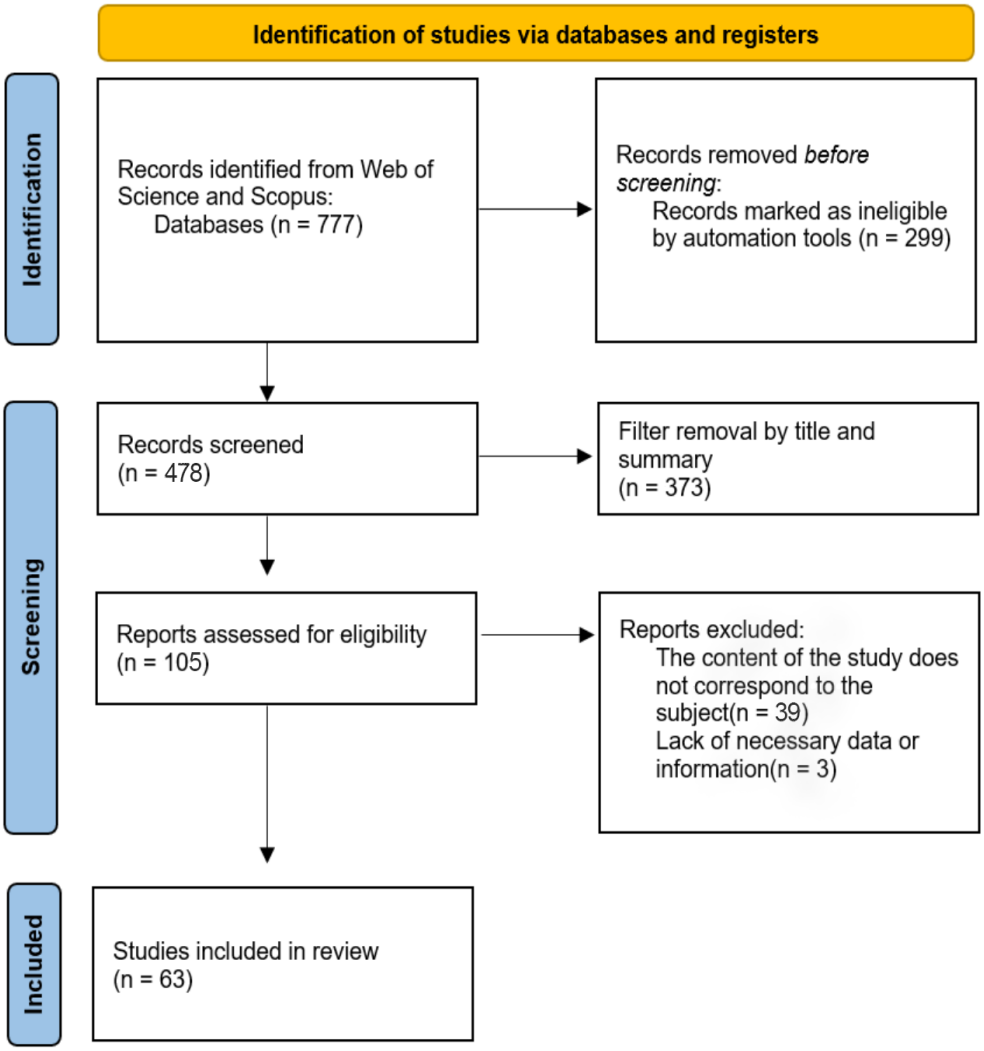


Figure 1. Study flowchart (PRISMA flowchart).

3.2. Preliminary Systematic Screening of Literature

Among the 777 initial results retrieved (Scopus: 461 articles, Web of Science: 316 articles), after removing duplicate publications (204 articles), review articles (30 articles), articles not meeting the required publication date (61 articles), and those for which the full text could not be obtained (4 articles), a total of 478 articles were preliminarily identified as meeting the search criteria.

3.3. Abstract Screening Based on NLP

We further screened the abstracts of the 478 articles using the spaCy library in Python 3.6 [56] to ensure that the literature met the requirements of the SPIDER framework. spaCy is a text preprocessing library in the field of natural language processing (NLP), offering functions such as tokenization, part-of-speech tagging, dependency parsing, lemmatization, and named entity recognition [57]. It is particularly suitable for large-scale semantic recognition in medium-sized texts [58].

We adapted the keyword list recognized by spaCy according to the SPIDER framework, as detailed below (Table 1).

Table 1. SPIDER framework keyword search table.

Criteria	Key Terms
Sample	Citizens, community, group, individuals, inhabitants, participants, pedestrians, people, population, public, residents, respondents, users, volunteers
Phenomenon of interest	accessibility, attention, awareness, behavior, comfort, decision, emotion, engagement, experience, feeling, interaction, perception, quality, reaction, response, safety, satisfaction, stress, usability, well-being
Design	approach, audit, case study, data, design, experiment, framework, method, model, observation, protocol, survey, system, test, tool
Evaluation	appraisal, assessment, comparison, effect, effectiveness, evaluation, findings, impact, measurement, outcome, performance, quality, rating, results, review, success, validation
Research type	analysis, comparative, descriptive, experimental, mixed method, quantitative, research, study, survey, trial

After NLP screening, we obtained 105 articles whose main research focus aligns completely with the SPIDER framework. These articles are considered to have statistical potential.

3.4. Feasibility Test

Among the 105 articles, two scholars in urban and regional planning served as independent reviewers to assess content usability. After they evaluated the usability of the search results and completed data extraction, coding, and thematic analysis, we conducted an inter-rater consistency check on their outputs. This approach was adopted to ensure scientific rigor and accuracy while minimizing individual bias. We used Cohen's Kappa method applied in binary classification tasks, to evaluate the consistency results (Appendix A) [59].

We believe that when consistency exists above 70%, it indicates that the two reviewers have similar intention trends. According to the calculation results, the Kappa value is 76.96%, indicating that the qualitative data of the statistics are available. Ultimately, we screened out 63 research articles that met the research requirements and had a complete SPIDER framework, and these were included in the data statistics.

3.5. Evaluation Framework

The built environment is diverse, and the differing needs of various populations result in environmental audits that also exhibit a diversity of characteristics [60]. To more comprehensively evaluate both qualitative and quantitative perception data in built environment audits, and to standardize the performance of AIGC in review and synthesis, it is necessary to design a detailed and rigorous evaluation system. For this purpose, we constructed targeted detailed textual descriptions under the SPIDER framework:

Basic Information: Year of publication, country of publication, research location.

S (Sample): Age group of the study population, gender.

PI (Phenomenon of Interest): Type of environment, spatial scale of the environment, environmental characteristics.

D (Design): Data sources, data collection methods, data collection frequency.

E (Evaluation): Data stream indicators, data analysis methods.

R (Research type): Type of study.

3.6. Information Extraction Based on LLMs

We utilized the official API interfaces provided by OpenAI and Anthropic to call the world's leading large language models—GPT-4, GPT-4-Optimized, and Claude-3.5-sonnet—to extract qualitative and quantitative information from 63 articles in three rounds, respectively. Using the pandas library, we stored the extracted results from each article in dictionary format and ultimately compiled them into a new Excel file. Based on information from the official websites and relevant literature [61], we summarized the differences among the three large language models in processing and synthesizing batches of literature, as shown in the table below (Table 2).

Table 2. Comparison of different LLM language models.

	Claude-3.5-Sonnet	**GPT-4-Optimized**	**GPT-4**
Release period	2024	2023	2023
Contextual logical reasoning	Excellent performance, suitable for long conversations and document analysis	Similar to GPT-4, the efficiency of long text processing is optimized	Strong, but the speed decreases as the length of the context increases
Language understanding	Good at understanding the context of discourse, especially in long academic papers to maintain consistency	Slightly less efficient than GPT-4, but more efficient for shallow tasks (such as quick summaries)	Good at understanding complex language expression and implicit semantics, dealing with fuzzy, ambiguous content
Multimodal processing	Good at explaining charts and graphs, able to extract text from imperfect images and other tasks	Supports analysis of image and table input analysis	Supports analysis of image and table input analysis
Computing costs	Input: 3 $/1 M tokens output: 15 $/1 M tokens	Input: 5 $/1 M tokens output: 15 $/1 M tokens	Input: 3 $/1 M Output: 6 $/1 M tokens
Prejudice and neutrality	Bias control is excellent, with a particular emphasis on the neutrality and safety of language in dialogue	Similar to GPT-4, some biases were reduced after optimization	There is some control, but it may take multiple optimizations to completely eliminate bias

3.7. Elimination of Potential Bias and Distortion

To avoid potential distortion and bias associated with AIGC technology, all AI-generated analytical results were subject to manual review and verification. However, this usability verification process is significantly faster than traditional manual synthesis. In this study, we manually conducted the following two checks and optimizations on the analytical data:

1. Cross-Validation of Model Results
   We compared the analytical and summary results of the three LLMs—GPT-4, GPT-4-Optimized, and Claude-3.5-sonnet. When the results of the three models were judged by the researchers to be homogeneous, they were considered as having passed the review. If there were differences among the results, they were identified as abnormal analytical values, and the relevant articles were subjected to manual review.
2. Manual Re-examination Validation
   We manually reviewed the results of AIGC analysis for the 63 articles against their original texts. To ensure the scientific rigor of the review and coding process, one reviewer first referred to the AI analysis results and annotated the corresponding information in the original article; another reviewer independently checked for in-

consistencies in the article content without reference to the AI results. To mitigate single-reviewer bias without altering the workflow, before making the final decision in cases of disagreement, the first author sought a brief opinion from another co-author as a reference and strictly adhered to the pre-specified inclusion/exclusion criteria. All coding processes were completed using MaxQDA 2024 software.

4. Results

4.1. Overview of Study and Global Distribution

According to this study's inclusion criteria, a total of 63 relevant articles published after 2005 were included for quantitative analysis (Figure 2a,b). These 63 studies span a wide range of spatial scales, including intercontinental, national, provincial, county, and street levels, and are distributed across 17 countries worldwide. Among them, Australia contributed the most studies (n = 29), followed by the United States (n = 10) and China (n = 5). Overall, over the past two decades, the number of studies employing perception data streams for built-environment audits has shown a steady upward trend, reflecting the field's ongoing development and growing academic attention.



Figure 2. Publication year and research country of the articles.

4.2. Built Environment Audit Objects

The studies included span five spatial scales—city, community, street, block, and building—and are analysed across seven environmental dimensions (Figure 3).

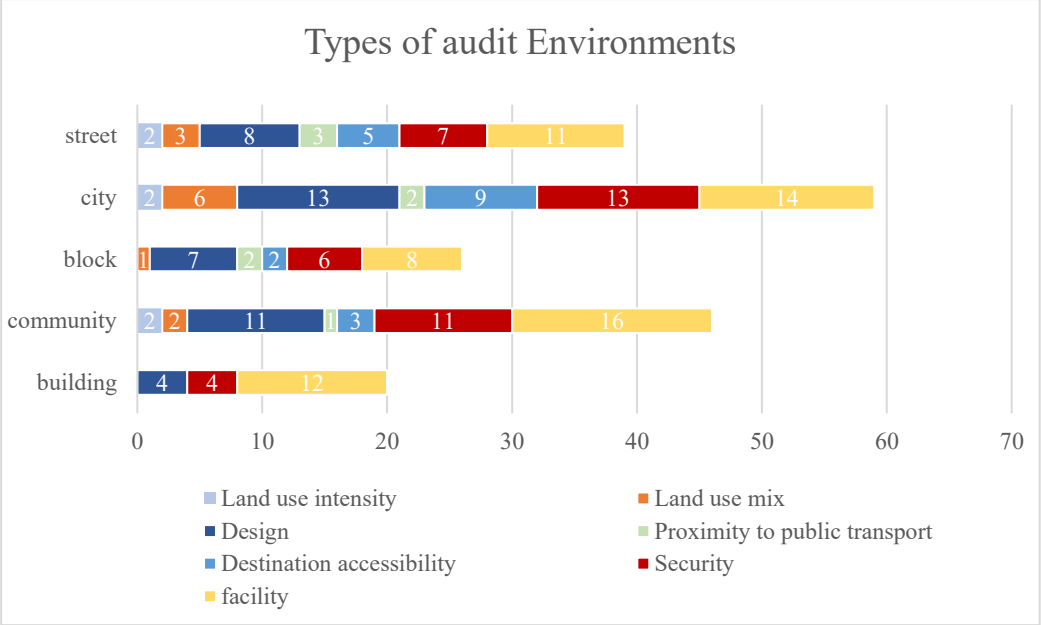


Figure 3. Types of audit environments.

The city scale is most frequently audited (49 studies), with research focusing on spatial structure, resource allocation, and factors such as design, safety, and facilities, underscoring their role in urban health. The community scale follows closely (46 studies), reflecting its importance for residents' daily life and social interactions. At this scale, facilities, design, and safety are key areas of focus, with additional attention to land use and accessibility.

Street and block audits address the usability and accessibility of the environment at the meso level. Street audits emphasize design and facility provision, while block audits offer a balanced assessment of multiple dimensions.

Building-scale audits are less common (20 studies), mainly focusing on design and facility configuration, highlighting the micro scale's relevance to individual health impacts.

Analysis of the seven environmental dimensions reveals that facilities are the most frequently assessed (53 occurrences) across all scales, followed by design (50) and safety (41). Land use intensity (9) and land use mix (15) are mainly addressed at the city and community levels. Destination accessibility (28) and public transit accessibility (7) are considered across several scales, with the latter being less common.

In summary, macro and meso scales (city, community, street, block) focus primarily on accessibility, resource support, and environmental quality, while the micro scale emphasizes spatial details and user experience. Facilities, design, and safety are universally important.

4.3. Categories and Analysis Methods of Perception Data Stream

4.3.1. Categories of Perception Data Stream

After synthesizing indicators of perception data streams from the existing literature, this paper classifies them into three categories: behavioral, psychological, and physiological, as shown in Table 3. Behavioral data primarily captures individuals' activity patterns in the built environment. Psychological data emphasizes subjective experiences and environmental appraisal, including satisfaction, well-being, affect, and perceived stress. Physiological

data, typically obtained via wearables or health-monitoring systems, index individuals' physiological state and anthropometrics.

Table 3. Categories of perception data stream.

Data Type	Specific Metrics	Type Description
Behavioral	Walking frequency, cycling time, sedentary time, path use frequency, physical activity participation frequency, number of pedestrians, etc.	The individual's clear activities and path choices in the built environment reflect the actual response and interaction mode of people to the environment.
Psychological	Environmental satisfaction, well-being, perceived stress, attachment, safety, recognition, subjective evaluation, etc.	The individual's subjective evaluation, emotional response and cognitive state of the surrounding environment reflect the psychological process of "the environment is perceived".
Physiological	Heart rate, blood pressure, BMI, blood tests, asthma rate, obesity indicators, etc.	Reveal the objective influence of the environment on physical health through physiological states and biological signals reflected by body sensors or health measurements.

4.3.2. Collection Tools of Perception Data Stream

A statistical analysis of the 63 studies shows that most employed 1–2 audit tools each. Based on their functions and scope, these tools are categorized as illustrated in Figure 4.

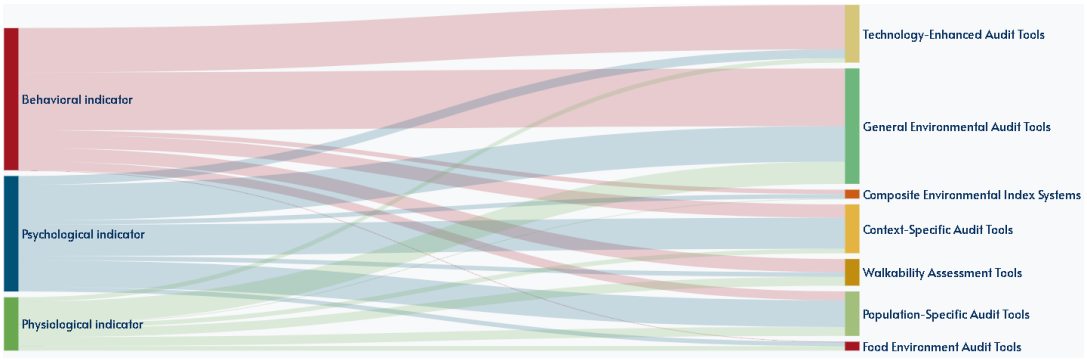


Figure 4. Categories and descriptions of audit tools.

The results demonstrate clear differences in audit tool suitability for various perception data types, reflecting the distinct functions of each tool category. Behavioral indicators, the most commonly used, are supported by general environmental audit tools, walkability assessment tools, and technology-enhanced audit tools. Psychological indicators, due to their subjective and contextual nature, are assessed with general, context-specific, and population-specific tools, enabling more nuanced evaluation of emotions and social perceptions. Physiological indicators are mainly paired with technology-enhanced and general audit tools, leveraging sensor data and computer vision for consistency.

General environmental audit tools are the most frequently employed (n = 26), demonstrating broad adaptability across different indicator types. Technology-enhanced tools follow (n = 13), utilizing street view imagery, and remote sensing to increase efficiency, though their generalizability may be limited outside their original training contexts.

Context-specific and population-specific tools are used 11 and 10 times, respectively, mainly for targeted studies in specific environments or groups, often combined for detailed analyses. Walkability assessment tools focus on behavioral and psychological aspects

related to physical activity. Composite environmental index systems and food environment audit tools are least used, serving primarily as supplementary instruments for policy and health assessments.

4.3.3. Analysis Approaches of Perception Data Stream

A variety of analytical methods are employed to process perception data streams and explore the relationships between environmental factors and health perceptions (Figure 5). Regression analysis is most used, particularly for behavioral data, with multiple linear and logistic models assessing the effects of built environment factors on outcomes like travel frequency or activity duration. Where possible, these models support causal inference; otherwise, they estimate associations. Comparative analyses are frequently applied to identify perceptual differences across groups, locations, or contexts, making them especially suitable for psychological and behavioral indicators. These analyses reveal social and spatial disparities and inform targeted interventions. Reliability and correlation analyses are more common in psychological and physiological research, assessing the stability of measurement tools and identifying preliminary relationships between variables. Although less frequently used, clustering and dimensionality reduction methods are valuable for managing high-dimensional perception data, aiding in category discovery and variable simplification. This reflects a trend toward more advanced and diverse analytical techniques.

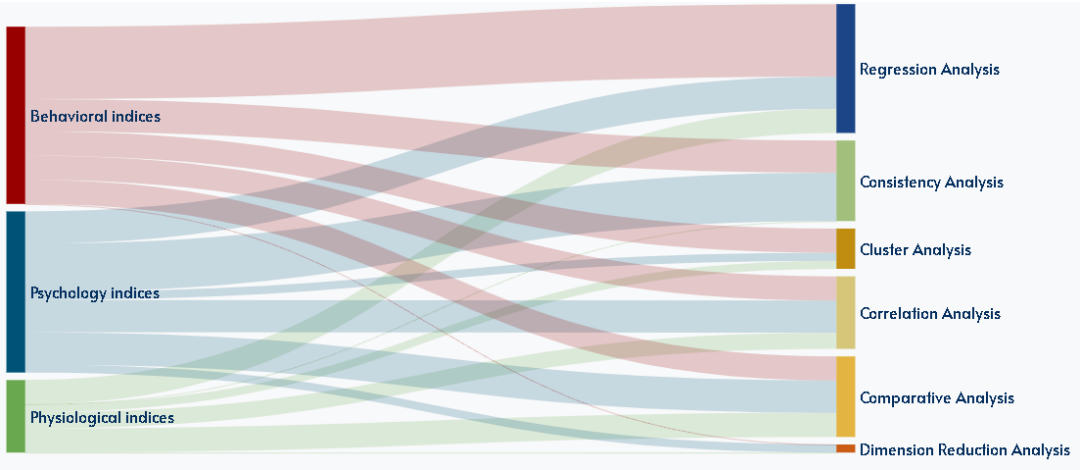


Figure 5. Categories and descriptions of analysis approaches.

In summary, analytical methods are chosen based on perception data type and research goals: behavioral data emphasize causal inference and comparisons, psychological data focus on measurement reliability and structure, and physiological data prioritize variable simplification and temporal association. Collectively, these approaches form a comprehensive analytical framework for studying environmental health perceptions.

4.4. Adaptability Analysis Between Perception Data Stream and the Built Environment

Analysis of perception data type distribution across spatial scales and built environment characteristics reveals clear patterns of adaptability among behavioral, psychological, and physiological indicators (Figure 6). Behavioral data are mainly used at larger spatial scales—community, city, and street—where they effectively capture activity patterns and travel behaviors in open environments but are less applicable at the building scale due to limited spatial or temporal granularity.

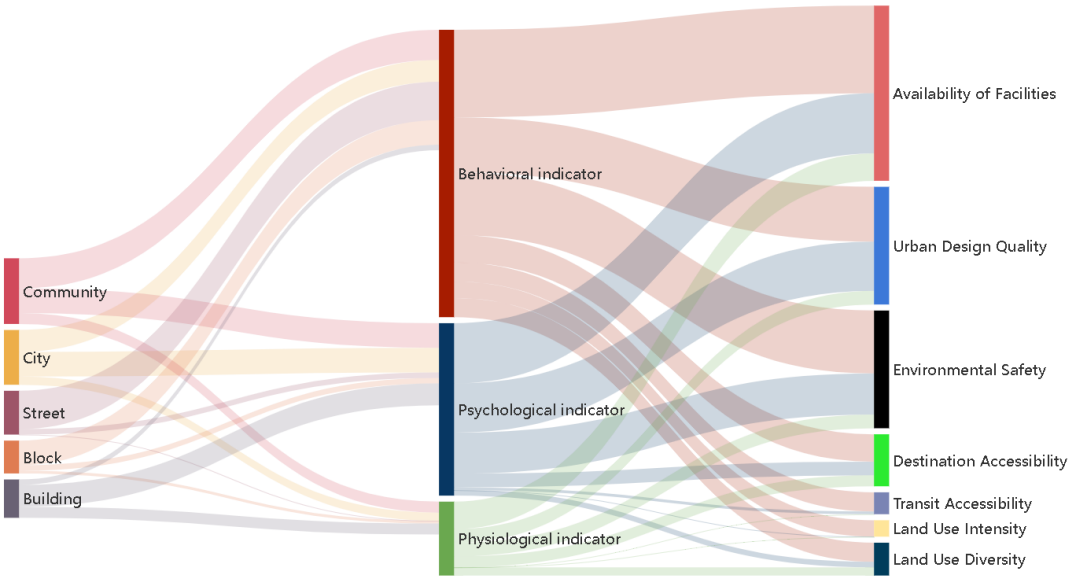


Figure 6. Perception data streams at different scales.

In contrast, psychological and physiological data show broader applicability across community, city, and building scales. They are well-suited for assessing emotional responses and physiological stress related to neighborhood safety, comfort, and facility access at the community level, as well as reactions to urban factors like spatial equity and noise at the city level. At the building scale, these data types capture the immediate effects of indoor environmental factors on occupant well-being.

Regarding built environmental features, behavioral indicators are central for structural elements like land use intensity, diversity, and transit accessibility. For experiential attributes—such as environmental safety and urban design quality—both behavioral and psychological data are key. Functional features like facility and destination accessibility show a more balanced use of all three data types, though behavioral data remain predominant. These patterns underscore the complementary roles of different perception data streams in the comprehensive assessment of built environments across various scales and features.

5. Discussion

5.1. Aligning Perception Data Types with Built Environment Audit

The adaptability of different perception data types to various spatial scales and audit elements is fundamentally determined by their intrinsic properties, their modes of information expression, levels of perception, and data collection mechanisms.

Behavioral data is best suited for larger spatial units such as cities, communities, and streets, as they capture observable movement patterns and activities within open and continuous environments. Macro-scale structures like road networks and land use patterns provide the necessary context for collecting and analyzing such data, making them effective for revealing population flows and spatial preferences at these scales. However, at the building scale, the enclosed environment and diversity of localized activities limit the representativeness and utility of behavioral data.

By contrast, psychological and physiological data are more appropriate for building and community scales due to their sensitivity to micro-environmental features and capacity for immediate response. Psychological data reflects subjective perceptions and emotions. While physiological data capture real-time bodily responses at finer spatial scales, variables like lighting, ventilation, noise, and privacy can directly influence psychological states and

physiological indicators, making these data types particularly valuable for assessing indoor and neighborhood environments.

When considering built environment audit elements, the reliance on different perception data types reflects the inherent attributes of the indicators. Structural features, such as land use intensity, diversity, and proximity to transit—emphasize efficiency and spatial allocation, thus favoring behavioral data for objective evaluation of usage patterns. In contrast, human-centered elements like environmental safety and urban design require psychological data to capture subjective experiences, supplemented by behavioral data to understand usage and avoidance patterns. For function-oriented elements such as facility and destination accessibility, all three data types are relevant: behavioral data document actual access behaviors, psychological data assess subjective evaluations of accessibility, and physiological data provide insights into physical stress and fatigue related to commuting or facility use.

Integrating perception data streams into built environment audits provides unique advantages for rigorously assessing environmental diversity and complexity. However, their application is constrained by the high cost and technical difficulty of data collection, as well as potential biases arising from the representativeness of test populations. Furthermore, the extent to which gender, ethnicity, and cultural factors shape physiological, psychological, and behavioral responses requires further validation.

5.2. Advantages of Using AIGC Techniques to Write Systematic Reviews

Compared with traditional review methods, AIGC leverages LLMs and NLP technologies, represented by models such as GPT-4, GPT-4o, Claude-3.5-sonnet, to achieve substantial optimization of the systematic literature analysis process in several key aspects. Firstly, in semantic retrieval and thematic classification, the powerful language understanding abilities of these models enable researchers to rapidly identify literature highly relevant to the research question and automatically recognize its themes, disciplinary background, and variable characteristics, significantly improving the efficiency of preliminary screening and classification. Secondly, AIGC can automatically extract key terms, variables, and indicators from the literature, facilitating the construction of summary tables that encompass all research subjects and content, thus enhancing the structural organization of the review. In addition, AIGC demonstrates a certain degree of auxiliary verification and error correction capability within systematic reviews. Specifically, models can quickly extract and structurally summarize large-scale literature content, generating preliminary analysis results such as the specific research subjects and frequency statistics of research variables. This automation process not only significantly reduces manual labor but also provides a clear reference framework for subsequent manual verification.

When researchers review AIGC-generated categorizations or statistical results of variables, any discrepancies, omissions, or misinterpretations compared to the original text can be quickly located and corrected. This workflow enhances information processing efficiency as well as sensitivity and responsiveness to potential errors, thereby improving the overall accuracy and systematicity of real-world literature assessment. Compared with the traditional approach of reading and comparing articles one by one, AIGC provides a "reference framework" for rapid cross-validation that not only reduces cognitive load but also helps to avoid omissions and biases in manual judgment, greatly enhancing the accuracy and reliability of systematic reviews.

5.3. Prospects and Limitations of AIGC Technology in Reviews

AIGC technologies have demonstrated substantial efficiency gains in systematic evaluations within the built environment, particularly in literature summarization, variable

extraction, and structured data representation. Nevertheless, current applications remain largely confined to relatively isolated tasks and face persistent challenges when deployed across the end-to-end review workflow, including uneven model capabilities and limited adaptability. In parallel, this study implemented a data lock for the audit-related literature on the built environment in October 2024, and publications appearing thereafter were not included. Given that our evidence base spans 2000 to 2024, this temporal boundary is unlikely to materially alter the principal structural conclusions; however, we acknowledge the timeliness constraint it introduces and explicitly disclose it to ensure transparency. At the same time, AIGC entails potential risks, including the introduction or amplification of bias and challenges to reproducibility due to model version drift and stochastic outputs—which we mitigated through dual human review, reporting inter-rater agreement, and documenting prompts, parameters, and model versions.

To address these limitations, future work should systematically evaluate model performance at each key stage of the review process, delineate functional boundaries and optimal use cases for different models, and develop coordination mechanisms tailored to the requirements of systematic reviews. Going forward, identifying and adopting models that are better aligned with the domain and task will be essential to further reduce these risks, alongside incremental updates conducted under a consistent methodological framework with periodic extensions of the search and analysis window to capture the latest developments and test the robustness of conclusions. Establishing such coordination, risk-mitigation, and update frameworks will not only broaden the application scope of AIGC in literature analysis but also support the development of intelligent, reusable platforms for systematic review writing, thereby substantially enhancing knowledge integration efficiency and research transparency in complex domains such as the built environment.

5.4. Future Prospects for Applying Perception Data in Built Environment Audits

This review systematically reflects the broad application and valuable experiences of perception data in built environment audits in recent years. Although certain limitations remain, it is reasonable to anticipate that perception data–based audits will have broad application prospects in the future. Against the background of climate change, human health status and subjective perception are increasingly becoming key criteria for evaluating built environment quality and adjusting construction strategies.

In the future, several critical factors will influence the application prospects of perception data in built environment audits. Advances in data collection and processing technologies will undoubtedly reduce the cost, threshold, and difficulty of perception data application, improve the accuracy of evaluation, and enable further diffusion of this approach. The diversification of perception data sources will also enhance its applicability; in addition to active collection, the integration of public health records and big online data may enrich the types of perception data and thereby broaden their scope. Investment and leadership from the public sector, especially municipal authorities—will make large-scale integration of perception data into built environment audits more feasible. These advances can be operationalized within scenario-based assessment frameworks, including comparing the suitability of strategies by combining fuzzy expert systems with self-organizing maps. In parallel, coupling perception layers with ecosystem service demand can guide the allocation of urban facilities and nature-based solutions, thereby strengthening evidence-based prioritization and equity-sensitive implementation. Of course, this will also require the establishment of stronger privacy protection strategies through legislation and oversight, as well as the mitigation of potential inequities during application.

6. Conclusions

This study systematically reviewed 63 research articles published before July 4, 2024, on the use of perception data in built environment audits, focusing on data types, collection methods, and spatial scales. Heart rate and emotion-related measures emerged as the most common indicators, effectively capturing psychological and physiological responses. At the spatial scale, research predominantly targets community and block levels, with particular attention to roads, green spaces, and residential areas. These variations highlight the close link between perception indicators and spatial functions. The review also demonstrated the practical value of AIGC in improving the efficiency and accuracy of literature screening, extraction, and synthesis, while noting its limitations as a supportive tool rather than a substitute for critical interpretation. Together, these findings strengthen methodological foundations and point to promising directions for further exploration:

First, in terms of methods and evaluation frameworks, we recommend further validation within scenario-based assessment systems and exploring the combined use of Fuzzy Expert Systems and Self-Organizing Maps to compare the suitability of different strategies. Second, regarding data and governance foundations, technological advances should reduce costs and barriers, diversify data sources, and enable compliant integration with public health and online data, while strengthening privacy protection and equity safeguards at the institutional level, with particular emphasis on municipal investment and coordination. Third, for practice translation and impact evaluation, perception data layers should be coupled with ecosystem service needs and the configuration of roads and green spaces to support evidence-based prioritization and resource allocation, alongside longitudinal tracking and sensitivity analyses to assess policy and project effectiveness. Overall, positioning perception data as supportive evidence, advancing under the dual drivers of technology and governance, and iterating through verifiable impact evaluation loops constitute the key pathway to realizing its long-term value.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/buildings15203684/s1, Figure S1: PRISMA 2020 Checklist [62].

Author Contributions: A.T.: Writing—original draft, Writing—review & editing, Validation, Methodology, Formal analysis, Conceptualization, Software; Z.Y.: Writing—original draft, Writing—review & editing, Conceptualization, Supervision, Project administration, Resources; W.O.: Writing—review & editing, Validation, Software, Methodology, Data curation. All authors have read and agreed to the published version of the manuscript.

Funding: National Natural Science Foundation of China (52278050), the Fundamental Research Funds for the Central Universities (2242025S30068), National Natural Science Foundation of China (52378047).

Data Availability Statement: Data will be made available on request.

Acknowledgments: During the preparation of this work, the authors used GPT-4, GPT-4-Optimized, and Claude-3.5-sonnet in order to innovatively establish a systematic review research framework process. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article.

Conflicts of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Construction of the Confusion Matrix

To calculate Cohen's Kappa in a binary classification task, it is first necessary to construct a 2×2 confusion matrix, which displays the agreement and disagreement between the two evaluators in the binary classification (Table A1).The binary "positive/negative"

classes are defined as follows: a positive case is a study that fully meets the SPIDER inclusion criteria, is situated in a built-environment audit context, uses perception data streams, and allows extraction of standardized variables; a negative case is a study that does not use perception data streams or does not allow extraction of the target audit variables and is therefore not included in the synthesis. For example, a field audit of street segments using wearable heart-rate and affect measures, analyzed with GIS at the street-segment scale, is classified as positive; a study that only reports image-classification accuracy without an audit application and without extractable variables is classified as negative.

Table A1. Confusion matrix result.

	Reviewer I: Positive	**Reviewer II: Negative**
Reviewer I: positive	a (N = 63)	b (N = 3)
Reviewer II: negative	c (N = 8)	d (N = 31)

Calculate the observed consistency P_O and expected random consistency P_e.

The observed consistency P_O is the proportion of two evaluators reaching an agreement in the classification task. This is Equation (A1):

$$P_O = \frac{a+d}{a+b+c+d} \tag{A1}$$

The expected random consistency P_e is calculated based on the classification probabilities of the positive and negative classes for each evaluator. This is Equation (A2):

$$P_e = \left(\frac{(a+b)\times(a+c)}{(a+b+c+d)^2}\right) + \left(\frac{(c+d)\times(b+d)}{(a+b+c+d)^2}\right) \tag{A2}$$

Calculate the Kappa value, this is Equation (A3):

$$k = \frac{P_O - P_e}{1 - P_e} \tag{A3}$$

The Kappa value ranges from −1 to 1: 1 indicates complete consistency, and 0 indicates consistency that is the same as random consistency. We believe that when consistency exists above 70%, it indicates that the two reviewers have similar intention trends. According to the calculation results, the Kappa value is 76.96%, indicating that the qualitative data of the statistics are available. Ultimately, we screened out 63 research literatures that met the research requirements and had a complete SPIDER framework, and these were included in the data statistics.

References

1. Perez, C. Unleashing a golden age after the financial collapse: Drawing lessons from history. *Environ. Innov. Soc. Transit.* **2013**, *6*, 9–23. [CrossRef]
2. Haase, D.; Haase, A.; Rink, D. Conceptualizing the nexus between urban shrinkage and ecosystem services. *Landsc. Urban Plan.* **2014**, *132*, 159–169. [CrossRef]
3. Deng, C.; Ma, J. Viewing urban decay from the sky: A multi-scale analysis of residential vacancy in a shrinking U.S. city. *Landsc. Urban Plan.* **2015**, *141*, 88–99. [CrossRef]
4. UN Habitat. *World Cities Report 2016: Urbanization and Development: Emerging Futures 2016*; UN: New York, NY, USA, 2016.
5. Hurlimann, A.; March, A.; Bush, J.; Moosavi, S.; Browne, G.R.; Warren-Myers, G. Climate change transformation in built environments – A policy instrument framework. *Urban Clim.* **2024**, *53*, 101771. [CrossRef]
6. Watts, N.; Amann, M.; Arnell, N.; Ayeb-Karlsson, S.; Belesova, K.; Berry, H.; Bouley, T.; Boykoff, M.; Byass, P.; Cai, W.; et al. The 2018 report of the Lancet Countdown on health and climate change: Shaping the health of nations for centuries to come. *Lancet* **2018**, *392*, 2479–2514. [CrossRef]

7. Seguin, R.A.; Lo, B.K.; Sriram, U.; Connor, L.M.; Totta, A. Development and testing of a community audit tool to assess rural built environments: Inventories for Community Health Assessment in Rural Towns. *Prev. Med. Rep.* **2017**, *7*, 169–175. [CrossRef]
8. Spudys, P.; Jurelionis, A.; Fokaides, P. Conducting smart energy audits of buildings with the use of building information modelling. *Energy Build.* **2023**, *285*, 112884. [CrossRef]
9. Long, N.; Fleming, K.; CaraDonna, C.; Mosiman, C. BuildingSync: A schema for commercial building energy audit data exchange. *Dev. Built Environ.* **2021**, *7*, 100054. [CrossRef]
10. Saelens, B.E.; Handy, S.L. Built environment correlates of walking: A review. *Med. Sci. Sports Exerc.* **2008**, *40*, S550–S566. [CrossRef]
11. Cresswell, I.; Murphey, H.T. Australia State of the Environment 2016: Biodiversity, Independent Report to the Australian Government Minister for the Environment and Energy ResearchGate. Available online: https://www.researchgate.net/publication/315045059_Australia_state_of_the_environment_2016_biodiversity_independent_report_to_the_Australian_Government_Minister_for_the_Environment_and_Energy (accessed on 2 July 2025).
12. Altomonte, S.; Allen, J.; Bluyssen, P.M.; Brager, G.; Heschong, L.; Loder, A.; Schiavon, S.; Veitch, J.A.; Wang, L.; Wargocki, P. Ten questions concerning well-being in the built environment. *Build. Environ.* **2020**, *180*, 106949. [CrossRef]
13. Lanza, K.; Oluyomi, A.; Durand, C.; Gabriel, K.P.; Knell, G.; Hoelscher, D.M.; Ranjit, N.; Salvo, D.; Walker, T.J.; Kohl, H.W. Transit environments for physical activity: Relationship between micro-scale built environment features surrounding light rail stations and ridership in Houston, Texas. *J. Transp. Health* **2020**, *19*, 100924. [CrossRef]
14. Francesconi, M.; Flouri, E.; Kirkbride, J.B. The role of the built environment in the trajectories of cognitive ability and mental health across early and middle childhood: Results from a street audit tool in a general-population birth cohort. *J. Environ. Psychol.* **2022**, *82*, 101847. [CrossRef]
15. He, L.; Páez, A.; Liu, D. Built environment and violent crime: An environmental audit approach using Google Street View. *Comput. Environ. Urban Syst.* **2017**, *66*, 83–95. [CrossRef]
16. Nangia, C.; Singh, D.P.; Ali, S. Built Environment and Crime Against Women: An Overview. In Proceedings of the 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 10–11 January 2019; pp. 636–641.
17. Pliakas, T.; Hawkesworth, S.; Silverwood, R.J.; Nanchahal, K.; Grundy, C.; Armstrong, B.; Casas, J.P.; Morris, R.W.; Wilkinson, P.; Lock, K. Optimising measurement of health-related characteristics of the built environment: Comparing data collected by foot-based street audits, virtual street audits and routine secondary data sources. *Health Place* **2017**, *43*, 75–84. [CrossRef]
18. Park, H.; Brown, C.D.; Pearson, A.L. A systematic review of audit tools for evaluating the quality of green spaces in mental health research. *Health Place* **2024**, *86*, 103185. [CrossRef]
19. Desjardins, E.; Higgins, C.D.; Scott, D.M.; Apatu, E.; Páez, A. Using environmental audits and photo-journeys to compare objective attributes and bicyclists' perceptions of bicycle routes. *J. Transp. Health* **2021**, *22*, 101092. [CrossRef]
20. Lee, S.; Lee, C.; Won Nam, J.; Vernez Moudon, A.; Mendoza, J.A. Street environments and crime around low-income and minority schools: Adopting an environmental audit tool to assess crime prevention through environmental design (CPTED). *Landsc. Urban Plan.* **2023**, *232*, 104676. [CrossRef]
21. Yang, H.; Peng, J.; Lu, Y.; Wang, J.; Yan, X. Nonlinear impact of built environment on people with disabilities' metro use behavior. *Appl. Geogr.* **2024**, *169*, 103323. [CrossRef]
22. Camatti, N.; di Tollo, G.; Gastaldi, F.; Camerin, F. Cultural heritage reuse applying fuzzy expert knowledge and machine learning: Venice's fortresses case study. *Reg. Stud. Reg. Sci.* **2025**, *12*, 225–251. [CrossRef]
23. Plascak, J.J.; Llanos, A.A.M.; Qin, B.; Chavali, L.; Lin, Y.; Pawlish, K.S.; Goldman, N.; Hong, C.-C.; Demissie, K.; Bandera, E.V. Visual cues of the built environment and perceived stress among a cohort of black breast cancer survivors. *Health Place* **2021**, *67*, 102498. [CrossRef]
24. Saito, Y.; Oguma, Y.; Inoue, S.; Breugelmans, R.; Kikuchi, H.; Oka, K.; Okada, S.; Takeda, N.; Cain, K.L.; Sallis, J.F. Inter-rater reliability of streetscape audits using online observations: Microscale Audit of Pedestrian Streetscapes (MAPS) global in Japan. *Prev. Med. Rep.* **2022**, *30*, 102043. [CrossRef]
25. Longato, D.; Cortinovis, C.; Balzan, M.; Geneletti, D. A method to prioritize and allocate nature-based solutions in urban areas based on ecosystem service demand. *Landsc. Urban Plan.* **2023**, *235*, 104743. [CrossRef]
26. Gehl, J. Cities for People. *Int. J. Sustain. High. Educ.* **2010**, *12*. [CrossRef]
27. Mahajan, S. Back Matter. In *The Art of Insight in Science and Engineering: Mastering Complexity*; MIT Press: Cambridge, MA, USA, 2014; p. 390.
28. Figueiredo, M.; Eloy, S.; Marques, S.; Dias, L. Older people perceptions on the built environment: A scoping review. *Appl. Ergon.* **2023**, *108*, 103951. [CrossRef]
29. Christoforou, R.; Lange, S.; Schweiker, M. Individual differences in the definitions of health and well-being and the underlying promotional effect of the built environment. *J. Build. Eng.* **2024**, *84*, 108560. [CrossRef]

30. Koohsari, M.J.; Yasunaga, A.; McCormack, G.R.; Shibata, A.; Ishii, K.; Nakaya, T.; Hanibuchi, T.; Nagai, Y.; Oka, K. Depression among middle-aged adults in Japan: The role of the built environment design. *Landsc. Urban Plan.* **2023**, *231*, 104651. [CrossRef]
31. Ji, Y.; Feng, X.; Zhao, H.; Xu, X. Study on the elderly's perception of microclimate and activity time in residential communities. *Build. Environ.* **2024**, *266*, 112125. [CrossRef]
32. Ren, M.; Zheng, P. Towards smart product-service systems 2.0: A retrospect and prospect. *Adv. Eng. Inform.* **2024**, *61*, 102466. [CrossRef]
33. Kazemi, M.H.; Alvanchi, A. Application of NLP-based models in automated detection of risky contract statements written in complex script system. *Expert Syst. Appl.* **2025**, *259*, 125296. [CrossRef]
34. Liu, L.; Sevtsuk, A. Clarity or confusion: A review of computer vision street attributes in urban studies and planning. *Cities* **2024**, *150*, 105022. [CrossRef]
35. Li, Z.; Ma, J.; Tan, Y.; Guo, C.; Li, X. Combining physical approaches with deep learning techniques for urban building energy modeling: A comprehensive review and future research prospects. *Build. Environ.* **2023**, *246*, 110960. [CrossRef]
36. Benjira, W.; Atigui, F.; Bucher, B.; Grim-Yefsah, M.; Travers, N. Automated mapping between SDG indicators and open data: An LLM-augmented knowledge graph approach. *Data Knowl. Eng.* **2025**, *156*, 102405. [CrossRef]
37. Chung, S.; Moon, S.; Kim, J.; Kim, J.; Lim, S.; Chi, S. Comparing natural language processing (NLP) applications in construction and computer science using preferred reporting items for systematic reviews (PRISMA). *Autom. Constr.* **2023**, *154*, 105020. [CrossRef]
38. Wang, L.; Zhou, X.; Liu, J.; Cheng, G. Automated layout generation from sites to flats using GAN and transfer learning. *Autom. Constr.* **2024**, *166*, 105668. [CrossRef]
39. Lin, H.; Jiang, X.; Deng, X.; Bian, Z.; Fang, C.; Zhu, Y. Comparing AIGC and traditional idea generation methods: Evaluating their impact on creativity in the product design ideation phase. *Think. Ski. Creat.* **2024**, *54*, 101649. [CrossRef]
40. Shuai, B. A rationale-augmented NLP framework to identify unilateral contractual change risk for construction projects. *Comput. Ind.* **2023**, *149*, 103940. [CrossRef]
41. Li, F.; Yang, Y. Impact of Artificial Intelligence–Generated Content Labels on Perceived Accuracy, Message Credibility, and Sharing Intentions for Misinformation: Web-Based, Randomized, Controlled Experiment. *JMIR Form. Res.* **2024**, *8*. [CrossRef]
42. Sun, Y.; Sheng, D.; Zhou, Z.; Wu, Y. AI hallucination: Towards a comprehensive classification of distorted information in artificial intelligence-generated content. *Humanit. Soc. Sci. Commun.* **2024**, *11*, 1278. [CrossRef]
43. Dwivedi, A.; Soni, R. Impacts of urban heat island effect on critical urban infrastructure: A review of studies published between 2012 and 2022. *Environ. Rev.* **2024**, *32*, 457–469. [CrossRef]
44. Moufid, O.; Praharaj, S.; Jarar Oulidi, H. Digital technologies in urban regeneration: A systematic review of literature. *J. Urban Manag.* **2025**, *14*, 264–278. [CrossRef]
45. Su, P.; Yan, Y.; Li, H.; Wu, H.; Liu, C.; Huang, W. Images and deep learning in human and urban infrastructure interactions pertinent to sustainable urban studies: Review and perspective. *Int. J. Appl. Earth Obs. Geoinf.* **2025**, *136*, 104352. [CrossRef]
46. Guo, D.; Chen, H.; Wu, R.; Wang, Y. AIGC challenges and opportunities related to public safety: A case study of ChatGPT. *J. Saf. Sci. Resil.* **2023**, *4*, 329–339. [CrossRef]
47. Whyte, W.H. The Social Life of Small Urban Spaces | Publications—Project for Public Spaces. 1980. Available online: https://www.pps.org/product/the-social-life-of-small-urban-spaces (accessed on 2 July 2025).
48. Di Rienzo, M.; Rizzo, F.; Parati, G.; Brambilla, G.; Ferratini, M.; Castiglioni, P. MagIC System: A New Textile-Based Wearable Device for Biological Signal Monitoring. Applicability in Daily Life and Clinical Setting. In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 17–18 January 2005; pp. 7167–7169.
49. Clifton, K.J.; Livi Smith, A.D.; Rodriguez, D. The development and testing of an audit for the pedestrian environment. *Landsc. Urban Plan.* **2007**, *80*, 95–110. [CrossRef]
50. Rainham, D.; Krewski, D.; McDowell, I.; Sawada, M.; Liekens, B. Development of a wearable global positioning system for place and health research. *Int. J. Health Geogr.* **2008**, *7*, 59. [CrossRef]
51. Boarnet, M.G.; Forsyth, A.; Day, K.; Oakes, J.M. The Street Level Built Environment and Physical Activity and Walking: Results of a Predictive Validity Study for the Irvine Minnesota Inventory. *Environ. Behav.* **2011**, *43*, 735–775. [CrossRef]
52. Strath, S.J.; Greenwald, M.J.; Isaacs, R.; Hart, T.L.; Lenz, E.K.; Dondzila, C.J.; Swartz, A.M. Measured and perceived environmental characteristics are related to accelerometer defined physical activity in older adults. *Int. J. Behav. Nutr. Phys. Act.* **2012**, *9*, 40. [CrossRef]
53. James, K.L.; Randall, N.P.; Haddaway, N.R. A methodology for systematic mapping in environmental sciences. *Environ. Evid.* **2016**, *5*, 7. [CrossRef]
54. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. Declaración PRISMA 2020: Una guía actualizada para la publicación de revisiones sistemáticas. *Rev. Española De Cardiol.* **2021**, *74*, 790–799. [CrossRef]

55. Cooke, A.; Smith, D.; Booth, A. Beyond PICO: The SPIDER tool for qualitative evidence synthesis. *Qual. Health Res.* **2012**, *22*, 1435–1443. [CrossRef]
56. Honnibal, M.; Montani, I.; Van Landeghem, S.; Boyd, A.; Peters, H. *spaCy: Industrial-strength Natural Language Processing in Python*; version 3.7.2; Zenodo: Geneva, Switzerland, 2020. [CrossRef]
57. Spring, R.; Johnson, M. The possibility of improving automated calculation of measures of lexical richness for EFL writing: A comparison of the LCA, NLTK and SpaCy tools. *System* **2022**, *106*, 102770. [CrossRef]
58. Malik, A.; Behera, D.K.; Hota, J.; Swain, A.R. Ensemble graph neural networks for fake news detection using user engagement and text features. *Results Eng.* **2024**, *24*, 103081. [CrossRef]
59. Pérez, J.; Díaz, J.; Garcia-Martin, J.; Tabuenca, B. Systematic literature reviews in software engineering—enhancement of the study selection process using Cohen's Kappa statistic. *J. Syst. Softw.* **2020**, *168*, 110657. [CrossRef]
60. Zallio, M.; Clarkson, P.J. Inclusion, diversity, equity and accessibility in the built environment: A study of architectural design practice. *Build. Environ.* **2021**, *206*, 108352. [CrossRef]
61. Saderi, D.; Mahmoud, R.S.G.; Bender, G.; Oladoyin, O.O.; Ilegbusi, P.H.; Rahgozar, A.; Roy, M.; Machado, M.; Senst, B.; Akpan, C.A.N.; et al. Peer Review of "Towards Evaluating the Diagnostic Ability of LLMs (Preprint). " *JMIRx Med.* **2024**, *5*, e69830. [CrossRef]
62. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* **2021**, *372*, n71. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.