Chapter

Generative AI as a Tool for Human-Centric Correctional Design

Çiğsem Yağmur Yüksel

Abstract

Correctional architecture is a complex design field in which security, human dignity, rehabilitation, and environmental quality must be addressed simultaneously. Emerging generative artificial intelligence (AI) and simulation technologies offer new possibilities for exploring these competing spatial and institutional priorities. This chapter examines how generative AI and related computational technologies can support human-centred correctional design while accounting for human rights, security requirements, and diverse institutional contexts. A systematic literature review was conducted using Web of Science (WoS) and Scopus, integrating research on correctional architecture, environmental psychology, generative AI, and building simulation. The findings show that human-centred correctional design requires context-sensitive rather than universally transferable approaches and that AI can support the generation of spatial alternatives, visualisation, and performance evaluation through digital twins and simulation. Based on the synthesis, a six-stage human–AI collaboration framework is proposed. This chapter concludes that AI should function as an evidence-based decision-support tool, bounded by human rights standards, contextual knowledge, and human oversight.

Keywords: generative AI, correctional architecture, human-centred design, digital twins, human–AI collaboration

1. Introduction

Correctional environments embody a fundamental architectural, institutional, and ethical tension. While prisons operate as spaces of confinement, security, institutional control, and surveillance, they must also safeguard basic human dignity, health, family contact, and opportunities for purposeful activity and eventual reintegration. Reconciling these potentially competing requirements within the same built environment makes correctional facilities one of the most ethically complex building types [1, 2]. For this reason, a prison should be understood as a multi-layered experiential environment in which different user groups engage with the institution through their own unique spatial conditions. For incarcerated individuals, everyday life, movement, privacy, social interaction, and access to activities are

directly structured by the built environment [3]. For correctional staff, healthcare professionals, and administrators, spatial organisation influences working conditions, stress, safety, supervision, and professional interaction [3]. For the families of incarcerated individuals, visitation environments constitute an important spatial interface among institutional confinement, continued social relationships, and eventual reintegration into society [4]. Correctional design can therefore be approached as a complex human–building interaction problem with legal, ethical, social, and operational implications.

Within this context, human-centred correctional design can be understood as an architectural approach that seeks to protect human dignity, support the well-being of incarcerated individuals and staff, enhance environmental quality, and facilitate rehabilitation and reintegration while maintaining necessary levels of safety and institutional security [1, 2]. This approach is closely related to normative and rehabilitative models of correctional practice, in which the conditions of confinement are intended, as far as possible, to support reintegration. Spatial qualities such as daylight, access to nature, privacy, acoustic comfort, opportunities for purposeful activity, social interaction, and appropriate relationships between correctional staff and incarcerated individuals therefore become relevant components of the correctional environment rather than secondary architectural amenities [2, 3].

Likewise, architecture itself should not be assumed to produce rehabilitation or behavioural change independently. Spatial conditions can support or constrain privacy, interaction, environmental comfort, institutional relationships, access to programmes, and everyday experience [1–4], but these conditions operate within broader correctional systems involving institutional practices, healthcare, education, social support, security, and rehabilitation programmes. The role of human-centred architecture is therefore better understood as creating environmental conditions that may support these wider institutional objectives rather than determining individual behavioural outcomes.

Against this background, generative artificial intelligence (AI) offers emerging possibilities as a design-support technology rather than an autonomous decision-maker. In architecture, generative AI is increasingly utilised for image generation [5], design knowledge organisation [6], option generation and early-stage ideation [6, 7], structured prompt-based workflows [8], human-centred design [9], and the evaluation of emerging generative design approaches [10]. Within correctional architecture, where design decisions must simultaneously balance security, privacy, human rights, rehabilitation, environmental performance, and operational needs, AI-supported workflows enable designers to generate and evaluate multiple spatial scenarios, visualise alternative environments, and merge design exploration with performance analysis prior to construction or renovation [7].

However, the use of generative AI in correctional environments also raises significant ethical concerns. The AI-generated solutions may reproduce biases embedded within datasets, privilege easily quantifiable performance criteria over lived experience, or reinforce surveillance-oriented spatial logics. In correctional settings, such limitations may have consequences extending beyond functional discomfort or aesthetic inadequacy, potentially affecting privacy, dignity, accessibility, safety, and fundamental rights [1]. Therefore, AI-supported correctional design requires explicit human oversight, transparent evaluation

criteria, and a rights-based framework within which technological outputs can be critically assessed.

Although research on human-centred correctional environments and AI-supported architectural design [5–10] are both expanding, their direct intersection remains comparatively underexplored. In particular, limited attention has been given to how generative AI, digital twin technologies (dynamic digital representations of physical buildings that integrate spatial, environmental, and operational data), environmental simulation, and participatory AI-supported workflows might be integrated with established correctional design principles while simultaneously accounting for security requirements, human-rights standards, institutional realities, and ethical constraints. Within this framework, AI is positioned as a decision-support mechanism within a broader evidence-based and human-centred correctional design process.

Accordingly, this chapter aims to examine the opportunities, limitations, and implementation conditions of generative AI and related computational technologies as design-support tools for human-centred correctional design. The principal research question guiding the chapter is:

'How can generative AI and related computational technologies support human-centred correctional design while remaining responsive to human rights, security requirements, institutional constraints, and diverse correctional contexts?'

To address this question, this chapter first describes the systematic literature search strategy used to establish its multidisciplinary evidence base. It then examines the strengths and limitations of current prison systems and situates contemporary correctional design within a broader international context. It pays particular attention to unit management (UM) models as an existing organisational and spatial approach that may provide a bridge between security requirements and human-centred correctional practice. Building on this foundation, this chapter examines how generative AI, digital twin technologies, environmental simulation, and participatory workflows may support correctional design through a six-stage human–AI collaboration framework. These technological possibilities are subsequently evaluated in relation to minimum international standards, including the United Nations Standard Minimum Rules for the Treatment of Prisoners (the Nelson Mandela Rules), as well as practical constraints such as overcrowding, budget limitations, technological infrastructure, privacy, and institutional capacity. Finally, this chapter develops context-sensitive recommendations for the responsible integration of AI into correctional architectural design.

2. Methodology: Systematic literature search strategy

A systematic literature search and qualitative synthesis were conducted using Web of Science (WoS) and Scopus to integrate research on correctional architecture, environmental psychology, human-centred design, and AI. The search covered 2015–2026 and was restricted to peer-reviewed journal articles (Document Type: Article). Three analytical domains were examined: generative AI in architectural design, correctional architecture and environmental psychology, and digital twin/simulation approaches relevant to human–building interaction.

Three Boolean queries structured the initial search. Generative AI and architectural design were examined using: ('Generative AI' OR 'GenAI' OR

'large language models' OR 'prompt engineering') AND ('architectural design' OR 'spatial configuration' OR 'spatial ambience') AND ('restorative' OR 'rehabilitative' OR 'well-being'), yielding 11 WoS and 7 Scopus records. Correctional architecture and environmental psychology were searched through: ('correctional facility' OR 'prison design' OR 'jail architecture' OR 'correctional architecture') AND ('environmental psychology' OR 'restorative environment' OR 'human-centric' OR 'Nelson Mandela Rules'), yielding three WoS and four Scopus records. Finally, ('Digital Twin' OR 'digital simulation' OR 'virtual model') AND ('occupant behavior' OR 'human–building interaction' OR 'conflict reduction' OR 'institutional stress') AND ('prison' OR 'correctional' OR 'institutional') returned no records in either database.

Because the third query produced no results, two complementary searches were conducted: ('Digital Twin' OR 'digital simulation') AND ('occupant behavior' OR 'human-building interaction' OR 'crowd simulation'), restricted to peer-reviewed Open Access articles, yielded 33 WoS and 29 Scopus records and provided transferable approaches to behavioural and spatial simulation. A second query, ('prison design' OR 'correctional architecture') AND ('institutional stress' OR 'conflict' OR 'well-being' OR 'spatial ambience'), yielded nine WoS and seven Scopus records. The initial null result was therefore interpreted as indicating a limited direct intersection between digital twin research and human-centred correctional environments within the specified search parameters.

Across these searches, 103 records were retrieved (56 WoS; 47 Scopus) and imported into Zotero. Duplicate or substantially overlapping records were merged or removed, after which titles and abstracts were screened for relevance. Studies were retained when they addressed human-centred or rehabilitative correctional design, spatial–psychological or behavioural relationships, generative AI in architectural design, or transferable digital twin and simulation approaches. Studies without a relevant architectural, spatial, human, or behavioural dimension were excluded. Following consolidation and screening, a core corpus of 76 publications was retained for qualitative synthesis. Because duplicate management and relevance screening were conducted concurrently in Zotero, separate exclusion counts were not retrospectively assigned.

The core corpus was thematically organised and qualitatively synthesised to examine relationships among correctional architecture, environmental psychology, human-centred design, and AI-supported design and simulation. Following editorial review, a supplementary targeted search was conducted for concepts requiring further development, particularly UM, normalisation-oriented correctional models, and international minimum standards, using combinations of terms including 'unit management', 'direct supervision', 'correctional', 'prison', and 'housing unit'. These supplementary sources were used to strengthen specific sections and were treated separately from the original 76-publication core corpus.

3. Current prison systems: Navigating security, rehabilitation, and institutional order

Correctional architecture is fundamentally shaped by the requirements of security, control, surveillance, and institutional order. Prisons must regulate access, movement, visibility, and interaction between different user groups while simultaneously accommodating healthcare, education, work, visitation, recreation,

and administrative functions [1, 2]. Circulation systems (the corridors, routes, and controlled pathways that organise movement through the facility), controlled thresholds, observation points, zoning, and physical barriers, therefore, serve essential operational and security functions. Conventional static security, including secure boundaries, controlled access, surveillance technologies, and institutional procedures, remains necessary, particularly in medium- and high-security settings. However, contemporary correctional practice also incorporates dynamic security, which complements physical security through communication, regular interaction, and professional relationships between correctional officers and incarcerated individuals [11]. From an architectural perspective, this distinction is important because spatial configuration can either facilitate or restrict visibility, interaction, and relational forms of security.

At the same time, conventional prison environments may produce spatial and environmental conditions that intensify institutional stress (the psychological and operational strain associated with confinement, restricted autonomy, surveillance, and demanding working conditions). Correctional architecture also regulates access to privacy, social interaction, environmental stimulation, and meaningful activity [3, 4]. Restricted movement, excessive surveillance, noise, insufficient daylight, limited privacy, and poor access to nature may negatively influence everyday experience [12–16]. These limitations become more pronounced under overcrowded conditions, where pressure on accommodation and shared facilities can restrict movement, reduce access to educational and recreational spaces, compromise privacy, and limit participation in rehabilitative programmes [17].

The consequences of these spatial conditions extend to both incarcerated individuals and correctional staff. Correctional personnel work within environments characterised by continuous vigilance, interpersonal tension, and occupational stress, while spatial organisational influences visibility, communication, perceived safety, and opportunities for interaction [4, 18–22]. For incarcerated individuals, access to educational spaces, workshops, healthcare, recreation, outdoor areas, and visitation facilities provides the spatial infrastructure through which many rehabilitative and social programmes operate [13, 23, 24].

Consequently, current prison systems reveal an ongoing tension between security, institutional efficiency, human dignity, environmental quality, and rehabilitation. Human-centred correctional design does not require the abandonment of security; instead, it seeks a more balanced relationship among physical control, relational security, privacy, purposeful activity, staff well-being, and rehabilitative opportunities [1–4]. The appropriate balance varies according to security level, incarcerated population, institutional resources, penal policy, and cultural context. As a result, correctional models developed in one setting cannot be treated as universally transferable. The following section examines how these tensions have been addressed and operationalised across different international correctional contexts, with particular attention to normalisation-oriented and rehabilitative models.

4. Global perspectives on contemporary prison design

Contemporary correctional design reflects different approaches to balancing security, normalisation, human dignity, and rehabilitation. Although

human-centred models increasingly emphasise environmental quality, purposeful activity, interpersonal relationships, and reduced unnecessary deprivation [3, 4], these principles cannot be translated into a single universal prison model. Their implementation depends on security requirements, institutional resources, population characteristics, penal philosophy, and cultural context.

The Norwegian correctional system, particularly Halden Prison, provides a prominent example of a normalisation-oriented approach within a high-security setting. At Halden, conventional static security is complemented by dynamic security, which emphasises communication and professional relationships between correctional officers and incarcerated individuals [11]. This model demonstrates that rehabilitation and human-centred conditions can co-exist with substantial security requirements. However, empirical research at Halden reveals that incarcerated individuals do not necessarily perceive correctional officers as rehabilitative or therapeutic actors [11]. Thus, supportive architectural and institutional conditions may facilitate rehabilitation but cannot independently determine interpersonal or behavioural outcomes.

Türkiye represents a different correctional context, characterised by diverse institutional types and security classifications, including high-security, closed, open, women's, and juvenile facilities [25]. Recent institutional policies increasingly emphasise physical environments that support education, psychosocial services, vocational training, healthcare, social activities, and rehabilitation alongside security [25]. Within this context, an empirical study involving 120 juvenile detainees across three closed juvenile detention facilities found significant relationships between psychosocial well-being and perceived environmental dimensions, including structure, privacy, emotional feedback, freedom, support, and social interaction [26]. Perceived safety was also associated with several of these dimensions, while structure, expressed through clear rules, routines, and predictability, emerged as an influential component of institutional experience.

Comparing these distinct correctional contexts demonstrates that human-centred correctional design should be based on adaptable principles rather than transferable architectural models. While the Norwegian example foregrounds normalisation and dynamic security, evidence from the Turkish juvenile facilities highlights the importance of balancing structure and predictability with privacy, appropriate freedom, support, and social interaction. This does not imply opposing national models; rather, it demonstrates that common objectives such as dignity, security, environmental quality, and rehabilitation must be interpreted according to specific institutional and population needs. Therefore, the challenge lies in establishing context-sensitive spatial and operational strategies.

5. UM in prisons: Integrating direct supervision, security, and rehabilitation

Unit management is a decentralised correctional management model that organises incarcerated populations into smaller housing units supported by multidisciplinary staff teams. Its defining characteristics include direct staff presence within or close to housing units, increased staff–incarcerated person contact, and greater delegation of decision-making compared with centralised prison management [27]. By bringing management and supervision closer to everyday

prison life, UM seeks to improve communication, institutional responsiveness, and safety.

From a human-centred perspective, the significance of UM lies in connecting security, daily management, interpersonal interaction, and rehabilitative activity at a smaller operational scale. Direct supervision associated with this approach is intended to reduce the physical and social distance between correctional officers and incarcerated individuals and to support more regular, constructive interaction [20, 27]. Research has associated these approaches with increased communication, staff involvement, and, in some contexts, improved institutional safety [27].

However, these benefits are not inherent to the spatial model itself. Direct-supervision environments do not consistently produce greater perceived safety or improved staff–incarcerated person relationships; some staff have reported feeling less safe or less empowered within such settings [27]. Translating human-centred design proposals into functioning operational practice depends on appropriate spatial and infrastructural conditions, adequate staffing, comprehensive training, management support, and staff participation in design and operational decisions [20]. Similarly, evaluations of UM programmes indicate that rehabilitative outcomes vary across interventions, demonstrating that unit-based organisation alone cannot be assumed to reduce misconduct or recidivism [21].

Unit management should be understood as a socio-spatial and organisational framework that considers security, staff presence, daily activity, communication, and rehabilitation as interconnected components. Its manageable operational scale provides a relevant basis for exploring how AI-supported tools could evaluate occupancy, circulation, environmental conditions, interaction patterns, and alternative spatial configurations. Ultimately, the model serves as a conceptual bridge between established human-centred correctional practices and the computational design and simulation approaches examined in the following section.

6. AI-enhanced human-centric correctional design

6.1 Generative AI as a design-support tool

Generative AI is increasingly being used in architecture for image generation [5], design knowledge organisation [6], prompt-based exploration [22], human-centred design reasoning [9], and the generation and evaluation of alternative design solutions [7, 10]. Rather than functioning as an autonomous designer, AI can be positioned as a design-support tool that expands the range of alternatives considered during early design stages while leaving problem definition, interpretation, and ethical judgement under human responsibility [7, 23]. Positioning AI as a design-support tool rather than an autonomous actor is particularly important in correctional architecture, where design decisions must reconcile security, privacy, rehabilitation, environmental quality, and human dignity under conditions of institutional power.

The effectiveness of AI-supported exploration depends substantially on how design intentions are translated into prompts and evaluation criteria. Structured prompt frameworks demonstrate that environmental qualities such as greenery, daylight, natural materials, and restorative atmosphere can be explicitly incorporated into generative processes [22], while large language models can

organise architectural knowledge through pre-defined patterns and human-centred reasoning [9]. In correctional design, a prompt prioritising maximum visibility and surveillance may produce fundamentally different spatial alternatives from one incorporating graduated privacy, daylight, purposeful activity, restorative qualities, and reintegration objectives. Text-to-image and multimodal systems can make such alternatives visible during early design stages [5], but their outputs must subsequently be evaluated against spatial, experiential, ethical, and performance criteria rather than being accepted on visual quality alone [7, 10].

6.2 A human–AI collaboration framework

For human-centred correctional design, AI-supported decision-making can be structured through a six-stage human–AI collaboration process (**Figure 1**). First, relevant stakeholders – including incarcerated individuals or people with lived experience, correctional staff, healthcare professionals, families, designers, administrators, and oversight bodies – identify spatial problems and expectations. Second, designers translate these concerns into explicit values, prompts, and design criteria. Third, generative AI produces multiple alternative scenarios rather than a single optimised solution. Fourth, these alternatives are evaluated through experiential knowledge, ethical considerations, architectural expertise, and building-performance criteria. Fifth, selected alternatives are tested using environmental simulations, digital twins, crowd modelling, and technical review. Finally, design decisions and the influence of AI are documented to preserve accountability.

This workflow expands the design space while maintaining direct human judgement over the interpretation and selection of AI-generated alternatives. Such stakeholder participation is essential because AI cannot independently define concepts such as dignity, rehabilitation, privacy, or acceptable security for different

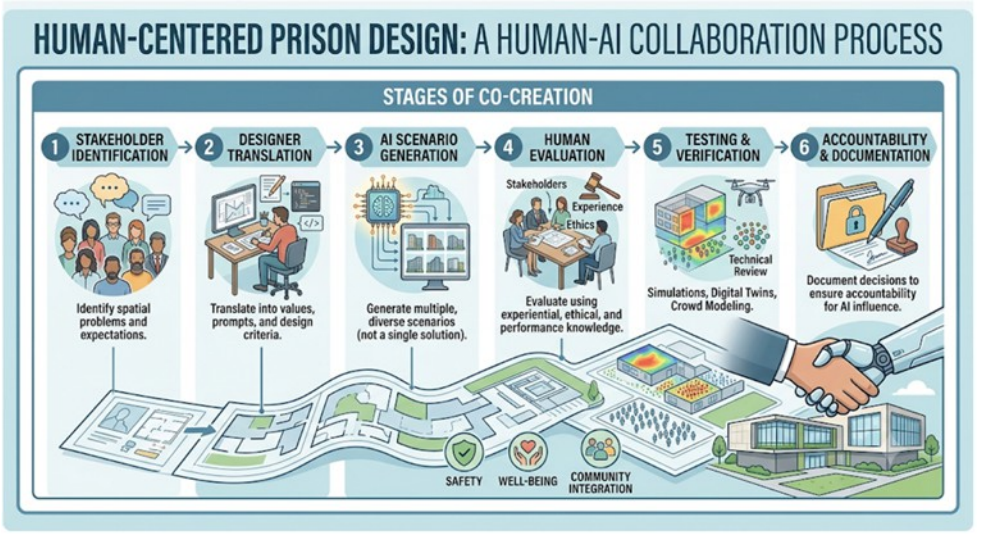


Figure 1.
A human–AI collaboration design process. Source: Prepared by the author using the Google image rendering model.

users. The AI-supported participatory environments can instead help non-expert stakeholders engage with architectural proposals by visualising alternative housing units, visitation areas, healthcare routes, outdoor spaces, or other institutional scenarios [24]. The AI-generated scenarios should be treated as provisional objects for discussion and revision, rather than final design decisions. Treating computational tools as reflective aids for participatory design aligns with evidence from correctional practice showing the importance of incorporating frontline and lived experience into spatial decision-making.

6.3 Digital twins, environmental simulation, and human–building interaction

Generative design becomes more analytically useful when combined with digital twins and environmental simulation. Digital twins conceptualise buildings as dynamic systems in which spatial, environmental, and occupant-related data can be evaluated together [28]. Occupant-centred approaches further demonstrate how users can participate in decisions concerning environmental comfort [29], while the Internet of Things (IoT), machine learning, and AI-supported systems enable the assessment of energy demand, HVAC performance, and building operation [30–33]. Applied cautiously to correctional environments, these technologies could enable alternative layouts to be assessed not only visually but also in relation to thermal comfort, ventilation, daylight, occupancy, circulation, and environmental performance.

A green approach to correctional architecture further extends human-centred design by integrating environmental sustainability with restorative spatial qualities. Access to daylight, natural ventilation, vegetation, outdoor spaces, and resource-efficient building systems can support both environmental performance and user well-being [12–16]. By uniting environmental performance with restorative design principles, AI-supported environmental simulation can assist in testing daylight, thermal comfort, ventilation, energy performance, and nature-integrated design alternatives, allowing sustainability objectives to be considered alongside security and rehabilitative requirements.

Crowd and movement simulations provide another relevant layer of evaluation. Correctional institutions involve highly regulated movement among housing units, healthcare, education, recreation, visitation, staff areas, and secure thresholds. Computational models can test collective movement, spatial density, congestion, and circulation performance during the architectural design process [34, 35]. Their relevance becomes particularly clear during emergencies, where locked doors and controlled circulation can conflict with rapid evacuation. Hybrid simulation methods can evaluate movement and life-safety risks before construction [36, 37]. Furthermore, scenario-based approaches developed for other complex buildings demonstrate how different patterns of use can be tested through simulated narratives [38]. In correctional design, such scenarios might include family visitation, access to healthcare, educational movement, outdoor exercise, staff-led de-escalation, or emergency evacuation.

6.4 Opportunities, ethical limits, and human oversight

The principal value of AI-supported correctional design lies in its capacity to combine scenario generation, visualisation, environmental analysis, and iterative

evaluation rather than in automated decision-making. Existing generative architectural research identifies rapid alternative generation, integration with performance criteria, and expansion of the design space as important opportunities [7, 10]. Integrating scenario generation, visualisation, and performance analysis may allow designers to compare different configurations before implementation and identify potential conflicts between environmental quality, circulation, security, and user needs.

However, the same technologies introduce significant ethical risks. The AI systems may reproduce historical biases present in their training data, generate visually convincing but contextually inappropriate solutions, or prioritise easily measured metrics such as spatial density or cost over the complex, lived experiences of occupants. Sensor networks and occupancy-monitoring tools introduce additional concerns: technologies designed to optimise environmental conditions (like heating, lighting, or air quality) can easily be repurposed to intensify institutional surveillance [34, 39, 40]. Research shows that human movement and spatial use can be continuously tracked through dense sensor networks; in correctional settings, therefore, a clear boundary must be established between monitoring building performance and tracking human behaviour. Principles of necessity, proportionality, transparency, privacy protection, and limited data use should explicitly govern AI-supported correctional design.

Artificial intelligence should function as a design-support mechanism within a human-led and evidence-based process. Its outputs may inform human decisions, but they do not replace institutional or ethical decision-making. It can expand design alternatives, visualise stakeholder concerns, and test spatial performance, but it cannot determine what constitutes justice, dignity, acceptable risk, or rehabilitation. Human judgement remains necessary to define these values, evaluate competing priorities, reject inappropriate AI-generated alternatives, and maintain accountability for final architectural decisions [9, 23]. Because these fundamental values cannot be automated, responsible AI-supported correctional design depends on integrating computational tools with participation, architectural expertise, empirical evidence, and explicit ethical oversight.

7. Minimum standards and international frameworks: The Nelson Mandela rules

Human-centred correctional design aligns spatial quality with the fundamental rights and well-being of occupants. Such a design orientation functions effectively only when embedded in an established rights-based normative framework, rather than relying solely on architectural preference or technological optimisation. The United Nations Standard Minimum Rules for the Treatment of Prisoners, commonly known as the Nelson Mandela Rules, establish fundamental principles concerning human dignity, humane treatment, health, accommodation, sanitation, exercise, healthcare, and rehabilitation [41]. From an architectural perspective, these principles translate into concrete spatial requirements related to adequate accommodation, ventilation, daylight, hygiene infrastructure, access to healthcare and activity spaces, and conditions that do not impose additional suffering beyond the deprivation of liberty itself. Existing research similarly demonstrates that prison architecture is directly connected to health, hygiene, nutrition, and institutional responsibility [15, 16].

The significance of these standards becomes particularly evident under conditions of overcrowding and resource scarcity. Research examining overcrowded prisons shows that insufficient sleeping space, limited access to sanitation, food, water, healthcare, and rehabilitative programmes can compromise both human dignity and institutional functioning [17]. Baffour et al. [17] explicitly relate such conditions to the Nelson Mandela Rules, noting that Rule 3 requires that imprisonment not aggravate the suffering inherent in the deprivation of liberty, while Rule 4 emphasises rehabilitation as part of reducing recidivism. These findings demonstrate that minimum standards are not abstract ethical aspirations; their implementation is directly shaped by spatial capacity, infrastructure, staffing, and access to essential services.

Under a rights-based framework, AI-supported design serves to assess compliance rather than redefine minimum rights. Generative AI and simulation tools assist designers in comparing alternative layouts, testing daylight and ventilation performance, identifying congestion and accessibility issues, estimating occupancy pressures, and evaluating circulation to healthcare, education, recreation, and visitation spaces [28–38]. However, compliance with minimum standards cannot be reduced to numerical optimisation; a technically efficient layout may still undermine dignity, privacy, social interaction, or equitable access to services. Therefore, the Nelson Mandela Rules serve as non-negotiable normative boundaries and evaluation criteria for assessing AI-generated alternatives.

Accordingly, the primary value of AI lies in detecting potential spatial conflicts before construction or renovation. Computational tools can rapidly identify where spatial capacity, daylight, ventilation, or circulation fail to meet established norms. However, determining whether a spatial environment is acceptable remains a human judgement. Architects, correctional staff, and human rights experts must retain full authority over these normative evaluations. Ultimately, responsible AI-supported design relies on a clear hierarchy: fundamental rights define the non-negotiable boundaries of space, while AI tests alternative ways to operate within them.

8. Challenges and implementation constraints

The implementation of AI-supported human-centred correctional design is often limited by physical, financial, and institutional constraints. Ageing infrastructure, limited space, and overcrowding place pressure on operational budgets, staffing, and access to rehabilitative programmes. Such physical limitations and resource pressures reduce the capacity of correctional institutions to implement spatial improvements or new technologies, even when human-centred design principles are formally supported. Research on overcrowded correctional facilities shows that spatial constraints and pressure on institutional resources can negatively affect infrastructure, essential services, and everyday living conditions [17]. Therefore, the feasibility of AI-supported design proposals should be assessed in relation to the available space, financial resources, staffing, and institutional capacity required for their implementation.

A second challenge concerns security levels, population characteristics, and institutional context. Minimum-, medium-, and maximum-security facilities require different balances among privacy, observation, movement, autonomy, and staff

presence, while juvenile, women's, healthcare-oriented, or other specialised populations may generate additional spatial and service requirements. The evidence reviewed in earlier sections indicates that even human-centred and relational models are not universally transferable; their effectiveness depends on staffing, organisational culture, management philosophy, and local penal conditions. In this context, AI should not be used to optimise a standardised 'ideal prison' but to compare context-sensitive alternatives against explicitly defined security, environmental, and rehabilitative criteria.

Population-specific needs require particular attention in AI-supported correctional design. Women, juveniles, older adults, persons with disabilities, and other vulnerable groups often present distinct requirements regarding privacy, healthcare, safety, accessibility, family contact, and rehabilitative programming. In this context, AI-generated design alternatives need to integrate gender- and population-sensitive criteria rather than rely on spatial assumptions derived from a standardised incarcerated population.

Technological implementation also introduces privacy, surveillance, bias, and accountability risks. Sensor networks, occupancy-monitoring systems, digital twins, and behavioural simulations may provide useful information for environmental performance and circulation analysis, but the same technologies can intensify institutional monitoring [34, 39, 40]. Generative systems may also reproduce dataset bias, produce contextually inappropriate design alternatives, or privilege variables that are easy to quantify over lived experience [7, 10]. Accordingly, data collection should be limited by principles of necessity, proportionality, transparency, and privacy protection, while AI-generated recommendations should remain subject to human evaluation and documented accountability.

Beyond technological capability, translating design proposals into actual operational practice depends on institutional acceptance and staff participation. Research on direct supervision demonstrates that progressive spatial or operational models can encounter resistance when frontline staff are excluded from design decisions, inadequately trained, or insufficiently supported [20]. For spatial and technological interventions to be truly effective in practice, correctional environments require adequate staffing, training, management commitment, interdisciplinary collaboration, and meaningful participation by those who live and work within these facilities. Such institutional realities reinforce the central argument of this chapter: AI can support correctional design only when technological innovation aligns with institutional capacity, human rights, and human judgement, rather than serving as a substitute for them.

9. Recommendations and conclusion

This chapter examined how generative AI and related computational technologies could support human-centred correctional design while remaining responsive to human rights, security requirements, institutional constraints, and diverse correctional contexts. Drawing on a systematic synthesis of literature across correctional architecture, environmental psychology, and AI-supported architectural design, this chapter's findings indicate that human-centred correctional design cannot be reduced to a single spatial model or technological solution. Rather, it requires a context-sensitive balance between security, dignity,

environmental quality, privacy, purposeful activity, relational interaction, and rehabilitation. The international and institutional perspectives examined in this chapter further demonstrate that such human-centred and evaluative design principles require adaptation to population characteristics, security levels, penal philosophy, institutional culture, and available resources. At the same time, the reviewed AI literature indicates that generative systems, digital twins, and simulation technologies can expand design alternatives, visualise spatial scenarios, and evaluate environmental and operational performance before architectural decisions are implemented [5–7, 9, 10, 22, 28–36, 38].

The AI-supported correctional design should be structured as a human-led, evidence-based, and context-sensitive decision-support process, rather than as an automated search for an 'ideal prison'. Generative systems can be used to produce and compare alternative spatial configurations [5–7, 9, 10, 22], while digital twins, environmental simulations, occupancy analysis, and movement modelling can evaluate the performance of these alternatives against criteria such as daylight, ventilation, accessibility, circulation, environmental comfort, security, privacy, and opportunities for purposeful activity [28–35]. Therefore, the proposed framework structures AI-supported design into six stages: stakeholder identification, translation of needs into design criteria, alternative generation, human evaluation, simulation and verification, and documented accountability.

For this human-led and context-sensitive approach to function effectively, AI-supported design processes need to operate within clear human-rights, ethical, and institutional boundaries. The Nelson Mandela Rules provide a normative baseline for evaluating proposed environments [41], while overcrowding, limited resources, existing infrastructure, staffing, and institutional capacity shape the practical feasibility of AI-supported interventions [17]. Human-centred principles also need to be adapted to context rather than treated as universally transferable, as security, privacy, autonomy, predictability, social interaction, and rehabilitation may require different spatial balances across populations and correctional settings. Furthermore, effective implementation requires the active involvement of relevant stakeholders and frontline staff, as their participation helps ensure that spatial and operational innovations respond to actual institutional needs and can be realistically applied in practice [20].

Accordingly, because the human-centred contribution of generative AI depends on how its capabilities are framed, evaluated, and governed, its role in correctional architecture is shaped by the framework within which it is applied. Its value lies not in replacing architects or correctional expertise, but in expanding the design space, testing potential consequences before implementation, and supporting more transparent and evidence-informed decision-making. Therefore, future AI-supported correctional design should integrate computational capabilities with human dignity, contextual knowledge, stakeholder participation, architectural expertise, and accountability. Such an approach allows technological innovation to support, rather than re-define, the fundamental human-centred objectives of correctional architecture.

Conflict of Interest

The author declares no conflict of interest.

Author details

Çiğsem Yağmur Yüksel
Çukurova University, Adana, Türkiye

*Address all correspondence to: cyyuksel@cu.edu.tr

IntechOpen

© 2026 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

References

[1] Bessemans C, Vandendriessche T. Ethics, architecture and prison design - a primer. The Journal of Architecture. 2024;**29**(7–8):967–990. DOI: 10.1080/13602365.2025.2490992

[2] Engstrom KV, Van Ginneken EFJC. Ethical prison architecture: A systematic literature review of prison design features related to wellbeing. Space and Culture. 2022;**25**(3):479–503. DOI: 10.1177/12063312221104211

[3] Mario B, Ricciardelli R, Johnston MS, Sibley MA. Examining the intersection of carceral space and well-being: Correctional officers' perspectives on old and new prison design. Canadian Geogr-Geogr Canadian. 2024;**68**(4):549–559. DOI: 10.1111/cag.12931

[4] Moran D, Jewkes Y. Linking the carceral and the punitive state: A review of research on prison architecture, design, technology and the lived experience of carceral space. Annales de Géographie. 2015;**2015**(702–703):163–184. DOI: 10.3917/ag.702.0163

[5] Ibrahim I, Abu Talib M, Ammar A, Aoul KAT, Abuimara T. Comparative and experimental analysis of leading text-to-image generative artificial intelligence models for regional residential architectural designs. Results in Engineering. 2026;**29**:108835. DOI: 10.1016/j.rineng.2025.108835

[6] Jang S, Roh H, Lee G. Generative AI in architectural design: Application, data, and evaluation methods. Automation in Construction. 2025;**174**:106174. DOI: 10.1016/j.autcon.2025.106174

[7] Zhuang X, Zhu P, Yang A, Caldas L. Machine learning for generative architectural design: Advancements, opportunities, and challenges. Automation in Construction. 2025;**174**:106129. DOI: 10.1016/j.autcon.2025.106129

[8] Malinowski A, Czarnul P, Czurylo K, Maciejewski M, Skowron P Multi-agent large-scale parallel crowd simulation. In: Koumoutsakos P, Lees M, Krzhizhanovskaya V, Dongarra J, Sloot P, editors. International Conference on Computational Science (iccs 2017). Amsterdam: Elsevier Science Bv; 2017. p. 917–926. (Procedia Computer Science). DOI: 10.1016/j.procs.2017.05.036

[9] Postle B, Salingaros NA. LLM and pattern language synthesis: A hybrid tool for human-centered architectural design. Buildings. 2025;**15**(14):2400. DOI: 10.3390/buildings15142400

[10] Yang Y, Li Y, Bai X, Zhang W, Chen S. Research progress and frontier trends in generative AI in architectural design. Buildings. 2026;**16**(2):388. DOI: 10.3390/buildings16020388

[11] Kilmer A, Abdel-Salam S, Silver IA. The uniform's in the way: Navigating the tension between security and therapeutic roles in a rehabilitation-focused prison in Norway. Criminal Justice and Behavior. 2023;**50**(4):521–540. DOI: 10.1177/00938548221143536

[12] Hopley R, Caulfield L, Jolly A. 'I'll live better, stay away from crime': Exploring the reintegration of former prisoners into the community through a music programme. Journal of Criminal

Psychology. 2023;**13**(4):351–366. DOI: 10.1108/JCP-12-2022-0033

[13] Murray C, Coyle B, Morgan H, Marder I, Woods D, Haughey T, et al. International consensus statement on the design, delivery and evaluation of sport-based interventions aimed at promoting social, psychological and physical well-being in prison. Bmj Open. 2024;**14**(1):e075792. DOI: 10.1136/bmjopen-2023-075792

[14] Stephenson T, Leaman J, O'Moore E, Tran A, Plugge E. Time out of cell and time in purposeful activity and adverse mental health outcomes amongst people in prison: A literature review. International Journal of Prisoner Health. 2021;**17**(1):54–68. DOI: 10.1108/IJPH-06-2020-0037

[15] Nomani MZM, Hussain Z. Health rights of prisoners and prison law reforms during COVID-19 pandemic in India. Bangladesh Journal of Medical Science. 2022;**21**(4):893–900. DOI: 10.3329/bjms.v21i4.60270

[16] Van Hout MC, Klankwarth UB, Glaum L, Stöver H. Violations of the right to food during deprivation of liberty: A global socio-legal assessment of United Nations human rights treaty bodies and the European Committee for the Prevention of Torture and Inhuman or Degrading Treatment or Punishment mission reporting on selected prisons since 2015. BMC Public Health. 2025;**25**(1):3818. DOI: 10.1186/s12889-025-25310-7

[17] Baffour FD, Francis AP, Chong MD, Harris N. Prison overcrowding and harsh conditions: Health and human rights concerns to persons in custody, staff, and the community. Criminal Justice and Behavior. 2024;**51**(3):375–400. DOI: 10.1177/00938548231219803

[18] Novek EM. That candy bar ain't free: Managing performances of masculinity in prison. Advances in Gender Research and. 2014;**18**(PARTB):195–223. DOI: 10.1108/S1529-21262014000018B012

[19] Kronfli N, Popovic N, Weiss E, Stoove M, Lafferty L, Price O, et al. Gatekeeping harm reduction in Canadian federal prisons: Perspectives on the threat risk assessment for the prison needle exchange program by prison administrative leadership. Addiction (Abingdon, England). 2026;**121**:2091–2103

[20] Sibley MA, Ricciardelli R, Johnston MS, Mario B. Direct supervision and the epistemic culture of prisons. Criminology & Criminal Justice. 2026;**26**(3):774–794. DOI: 10.1177/17488958241298486

[21] Lugo M, Wooldredge J, Pompoco A, Sullivan C, Latessa EJ. Assessing the impact of unit management programs on institutional misconduct and prison "returns". Justice Quarterly. 2019;**36**(1):59–92. DOI: 10.1080/07418825.2017.1357741

[22] Lee EJ, Park SJ. A structured prompt framework for AI-generated biophilic architectural spaces. Journal of Building Engineering. 2025;**111**:113326. DOI: 10.1016/j.jobe.2025.113326

[23] Hikmet S, Ozay N. Human-AI collaboration in architectural design education: Towards a conceptual framework. Buildings. 2026;**16**(6):1097. DOI: 10.3390/buildings16061097

[24] An H, Lu W, Wu L, Peng Z, Lou J. Meta-interaction: Deployable

framework integrating the metaverse and generative AI for participatory building design. Automation in Construction. 2025;**169**:105893. DOI: 10.1016/j.autcon.2024.105893

[25] T.C. Adalet Bakanlığı Adli Sicil ve İstatistik Genel Müdürlüğü [Republic of Türkiye, Ministry of Justice, General Directorate of Judicial Records and Statistics]. Adalet istatistikleri 2025 [Justice statistics 2025]. Ankara: T.C. Adalet Bakanlığı; 20Apr 2026. Available from: https://adlisicil.adalet.gov.tr/Resimler/SayfaDokuman/202604060901356772025Adalet%C4%B0statistikleriKitab%C4%B1.pdf [Accessed: 2026-November-14]

[26] Yüksel ÇY, Yeğin M. Psychosocial experiences of juvenile detainees in Türkiye: An analysis of environmental factors from an ecological systems perspective. International Journal of Offender Therapy and Comparative Criminology. 2026;**70**(5):561–582. DOI: 10.1177/0306624X251395224

[27] Farmer JF. Testing the conceptual path to correctional staff safety: A study of the implementation of unit management in two medium security state institutions in the USA. International Journal of the Criminal Justice Sciences. 2012;**7**(1):431–449

[28] Gnecco VM, Vittori F, Pisello AL. Digital twins for decoding human-building interaction in multi-domain test-rooms for environmental comfort and energy saving via graph representation. Energy and Buildings. 2023;**279**:112652. DOI: 10.1016/j.enbuild.2022.112652

[29] Saadatifar S, Sawyer AO, Byrne D. Occupant-centric digital twin: A case study on occupant engagement in thermal comfort decision-making. Archit-Switz. 2024;**4**(2):390–415. DOI: 10.3390/architecture4020022

[30] Abidin AZ, Enriko IKA, Pramudita AA. Leveraging IoT, digital twin and machine learning for smart energy audit in office building: A systematic literature review and recommendations. E-Prime - Advances in Electrical Engineering Electronic Energy. 2025;**14**:101124. DOI: 10.1016/j.prime.2025.101124

[31] Lu S, Zhou S, Ding Y, Kim MK, Yang B, Tian Z, et al. Exploring the comprehensive integration of artificial intelligence in optimizing HVAC system operations: A review and future outlook. Results in Engineering. 2025;**25**:103765. DOI: 10.1016/j.rineng.2024.103765

[32] Hosamo H, Mazzetto S. Performance evaluation of machine learning models for predicting energy consumption and occupant dissatisfaction in buildings. Buildings. 2025;**15**(1):39. DOI: 10.3390/buildings15010039

[33] Vittori F, Tan CF, Pisello AL, Chong A, Piselli C, Miller C. BIM-to-BRICK: Using graph modeling for IoT/BMS and spatial semantic data interoperability within digital data models of buildings. Energy and Buildings. 2025;**348**:116368. DOI: 10.1016/j.enbuild.2025.116368

[34] Jeong J, Park WK. Privacy-preserving labeling-free occupancy counting sensor based on ToF camera and clustering. ETRI Journal. 2025;**47**(5):841–854. DOI: 10.4218/etrij.2025-0022

[35] Liang H, Sun J, Wong SC, Liu X, Lam CC, Wong MO. Toward interactive crowd dynamics in virtual built

environment: A BIM-driven and macro-micro integrated crowd simulation framework. Developments in the Built Environment. 2026;**26**:100937. DOI: 10.1016/j.dibe.2026.100937

[36] DDACJ J, Wimalaratne SPW, Karunananda AS. Modelling goal selection of characters in primary groups in crowd simulations. International Journal of Simulation Modelling. 2016;**15**(4):585–596. DOI: 10.2507/IJSIMM15(4)1.323

[37] Kremyzas A, Jaklin N, Geraerts R. Towards social behavior in virtual-agent navigation. Science China Information Sciences. 2016;**59**(11):112102. DOI: 10.1007/s11432-016-0074-9

[38] Schaumann D, Breslav S, Goldstein R, Khan A, Kalay YE. Simulating use scenarios in hospitals using multi-agent narratives. Journal of Building Performance Simulation. 2017;**10**(5–6):636–652. DOI: 10.1080/19401493.2017.1332687

[39] Grubel J, Thrash T, Aguilar L, Gath-Morad M, Helal D, Sumner RW, et al. Dense indoor sensor networks: Towards passively sensing human presence with LoRAWAN. Pervasive and Mobile Computing. 2022;**84**:101640. DOI: 10.1016/j.pmcj.2022.101640

[40] Li B, Fitzgerald J, Schultz C. Modelling the impacts of crowds on occupants in the built environment-a static, rule-based approach to human perception and movement. Advanced Engineering Informatics. 2022;**51**:101452. DOI: 10.1016/j.aei.2021.101452

[41] United Nations. The United Nations Standard Minimum Rules for the Treatment of Prisoners. Vienna: United Nations Office on Drugs and Crime; 2015. Report V.16-00193