پنجمین کنفرانس
دانشجویان مهندسی کامپیوتر و فناوری اطلاعات و ارتباطات
5th Conference on
Computer Engineering, Information Technology and Communications Students

Generative Models and the Challenge of Representing Vernacular Architecture in Hot-Arid Climates

Negar Shahiddokht

Ph.D. in Artificial Intelligence, Amirkabir University of Technology, Tehran, Iran

ABSTRACT

Generative layout models are proposed as general instruments of architectural design, yet every large public floor-plan corpus on which they are trained originates in a narrow band of temperate and continental climates and in an extroverted, room-hub dwelling typology. This paper asks what follows when such a model is applied to hot-arid contexts whose historical typology is the introverted courtyard house. The contribution is addressed to the learning system rather than to the building: what a plan representation is able to encode, what a conditioning mechanism is able to constrain, and what a corpus-referenced metric is able to reward. The courtyard house is the test case, chosen because it makes those three limits measurable. A documentary audit of four principal corpora, together exceeding two hundred thousand annotated plans, establishes that none documents an origin in a hot-arid or semi-arid Köppen zone and that no unroofed interior court appears as a distinct spatial class in any of their published label vocabularies; two corpora released since, one of them explicitly assembled for geographical diversity, are shown not to alter the finding. We then argue that the resulting limitation is not a data shortage but a failure of encoding. Of five requirements of the typology, only one is a shortage of examples. The courtyard is a void that organises the plan yet carries no enclosure attribute; the constraint that habitable rooms do not open to the street cannot be written where the exterior is not a node; the privacy gradient is a property of topological depth that adjacency conditioning does not express; and seasonal migration between winter and summer quarters is absent from a static plan representation altogether. A model can therefore be enlarged with local data and remain structurally blind to what makes the typology work. We introduce the Corpus–Context Distance protocol, a five-dimension diagnostic applied before deployment that yields a disclosure statement for the design record, and we specify what a corpus adequate to hot-arid architecture would have to encode. The study is documentary and conceptual; its propositions are stated so that they can be refuted.

Keywords: *generative design, floor plan generation, training corpora, courtyard house, hot-arid climate, spatial privacy, algorithmic bias*

1. INTRODUCTION

Automated floor-plan generation has become one of the most productive research fronts at the intersection of machine learning and architecture. Within a few years the field moved from rule-based space planning to learned generation conditioned on a boundary [1], on an adjacency graph [2], [3], on a layout graph [4], and most recently on natural language [5]. The outputs are frequently indistinguishable from human-authored plans in user studies, and the tools built on them are marketed without geographic qualification: they generate plans, not plans for somewhere.

This paper takes that unqualified claim seriously and asks what it rests on. Every one of these models learns its notion of a plausible dwelling from a corpus, and the corpora are few. Four datasets recur across the majority of published work, and their provenance is documented: RPLAN comprises 80,788 plans of real residential buildings in Asia, assembled through a Chinese interior-design platform [1]; the layout research initiated by House-GAN drew 117,587 plans from a Japanese housing database, vectorised with an existing floor-plan transformation algorithm [2], [22]; CubiCasa5K contains 5,000 annotated Finnish plans [6]; and Modified Swiss Dwellings supplies 5,372 Swiss plans covering some 18,900 apartments [7], [23]. These are the reference points against which realism is measured and progress is claimed.

They share more than convenience. All four originate in temperate or continental climates. All four document extroverted apartment or single-family plans organised around an interior circulation core or a living room. None documents an origin in a hot-arid region, and in none of them is an unroofed interior court a distinct spatial class. A model trained on them has learned, with great precision, what a dwelling is in urban China, Japan, Finland and Switzerland.

The question this paper pursues is what happens when that model is asked to produce a dwelling for Yazd, Ghadames or Jaisalmer — places whose historical housing is organised around a courtyard, whose plans are introverted rather than extroverted, and whose spatial order encodes a social gradient that the corpus typology does not possess. The intuitive answer is that the model will underperform and that more data will fix it. We argue that this answer is wrong in an instructive way.

1.1 Position Relative to Existing Work

That generative systems carry cultural bias is not a new claim. Analyses of image generation trace bias in generated art to the composition of the training corpora and the art-historical canon they encode [8], and empirical testing of text-to-image systems reports markedly degraded output when African subjects are requested rather than unmarked ones [9]. The general mechanism by which an unexamined corpus becomes an unexamined norm is well established in the language-model literature [21]. Within architecture, narrative reviews of the technology's entry into practice stress how little of the situated and experiential dimension of space these systems capture [10]. The concern is registered inside the floor-plan literature as well: the authors of a recent multimodal floorplan dataset open by observing that existing work is confined to a single semantic category and region, typically apartments from one country [24]. This paper is complementary to that literature and differs from it in three respects. First, it is specific: instead of bias in general it examines one typology, one climate family and one technical task, which makes the claims checkable. Second, it is operational: rather than concluding that more inclusive data is needed, it asks whether inclusion is achievable within the present representation at all, and offers a diagnostic that practitioners can apply before deployment. Third, where the diversity critique is normally answered by broadening collection, Section 4.1 examines the two most recent attempts to do so and shows why breadth of collection does not by itself supply the missing categories.

1.2 Contributions

- *A documentary audit of the four principal public floor-plan corpora against climate zone, dwelling typology and room-label schema, extended to the two most recent large releases (Section 4).*
- *A specification of five requirements of hot-arid courtyard housing, drawn from the vernacular performance literature, and an assessment of whether each is representable in current plan encodings (Section 5).*
- *An argument that four of the five fail at the level of representation rather than data volume — three purely so, and one in representation and evaluation together — so that corpus enlargement alone cannot resolve them (Section 6).*
- *The Corpus–Context Distance protocol, a five-dimension pre-deployment diagnostic producing a disclosure statement for the design record (Section 7).*

1.3 Method and Its Limits

The study is documentary and conceptual. Corpus properties are taken from the publications that introduced the datasets and from the papers that report their use; typological requirements are taken from the peer-reviewed literature on courtyard performance in hot-arid regions. No model was trained, no plan was generated for this study, and no field measurement was undertaken. The claims should be read accordingly: they concern what published corpora and published encodings do and do not contain, which is verifiable from the record, and they generate predictions about model behaviour that Section 9 states in testable form. Coverage is limited to English-language sources and to publicly documented datasets; proprietary corpora held by software vendors are outside the audit and may differ.

2. BACKGROUND: LEARNED LAYOUT GENERATION

Automated space planning has a long history in rule-based and optimisation form, but the methods at issue here differ in that they induce a notion of plausibility from examples rather than executing a criterion an architect wrote. Three conditioning mechanisms dominate the recent literature, and each defines what a designer is able to ask for.

2.1 Conditioning Mechanisms

Boundary conditioning takes a site or unit outline and fills it. The influential formulation of this approach positions a living room first and then generates the remaining rooms iteratively, before resolving walls through an encoder-decoder network and refining them to vector form [1]. Graph conditioning takes a bubble diagram of typed rooms and their adjacencies and produces room geometry consistent with it, an approach introduced with relational generative adversarial networks [2] and extended through iterative refinement [3]; a related line conditions on a layout graph retrieved from a corpus of comparable plans [4]. Language conditioning replaces the diagram with a natural-language brief, supported by datasets that pair textual instructions with existing plans [5]. Diffusion and multimodal transformer methods have more recently been applied to the same task.

Whatever the mechanism, the pipeline has a fixed shape: a corpus supplies the notion of a plausible plan, a conditioning signal narrows it, and a metric certifies the result. The corpus is the only place in that pipeline where an assumption about how people live can enter, and it is the only element the user of the tool never sees.

2.2 How the Field Measures Success

The evaluation conventions were established alongside the methods. Graph-conditioned work proposed three measures — realism, diversity, and compatibility with the input constraint — with realism assessed against the distribution of plans in the training corpus [2]. Boundary-conditioned work relied on formative user studies comparing generated plans against human-authored ones [1]. A systematic review of 161 journal papers on generative artificial intelligence in architectural design confirms the pattern at scale: comparative evaluation is the dominant protocol, supported by subjective assessment frequently performed by the authors themselves, and image data dominates both input and output [20]. Section 6.4 returns to the consequences of anchoring realism to a corpus.

3. THE COURTYARD HOUSE AS A PERFORMANCE SYSTEM

Before assessing what corpora omit it is necessary to state precisely what the omitted typology does, because the courtyard house is frequently treated as a stylistic preference when it is in fact a coupled climatic and social instrument. The account below draws on the Iranian literature, which documents this typology in unusual quantitative detail; the courtyard house is of course far more widely distributed, and Section 8.2 considers how far the argument travels.

3.1 Climatic Function

Field studies of traditional houses in the hot-arid and semi-arid regions of Iran establish that the central courtyard operates as a microclimate modifier rather than as leftover space. Empirical work across fourteen traditional houses in five historical cities of the semi-arid (BS) zone found that orientation, dimensions and proportions of the courtyard, together with the ratio of opaque to transparent surfaces enclosing it, were resolved with sufficient consistency to constitute a design system rather than a convention [11]. A parallel survey of nine houses in the hot-arid BWhs mesoclimate reports the same variables converging on a comparable model [14], so the finding is not an artefact of one climate class. The mechanisms are well documented: shading of the enclosed void through the day, night-sky radiative cooling of exposed surfaces, stack-driven convective movement between the courtyard and the rooms, thermal storage in massive envelopes, and evaporative cooling from shallow pools and planting [12], [13].

Two features of this system matter for what follows. It is geometric: performance depends on the proportions of the void relative to the enclosing mass, not on the presence of a void as such. And it is seasonal: the same house is *inhabited differently in summer and winter, with rooms on opposing sides of the courtyard used at different times of year, so that the plan is a description of two dwellings sharing one envelope.*

3.2 Social Function

The courtyard also organises a social gradient. Space-syntax analysis of traditional Iranian houses shows a configuration in which yards and the corridors linking the outer and inner domains carry the highest connectivity while rooms and service spaces sit at maximal depth, a pattern that inverts in modern apartment plans where integration is distributed far more evenly [15]. The analytical apparatus is that of configurational analysis, in which a plan is read as a graph of permeability and social relations are recovered from its structure [18]. The gradient is realised through a sequence — an entrance portal, a vestibule that turns the line of sight, a corridor, then the courtyard — that filters access rather than merely providing it, so that a visitor received in the outer domain never obtains a view into the inner one [16]. The distinction between the outer and inner domains is long-standing in the stylistic literature on Iranian architecture, where it is treated as a property of the inward-facing plan rather than as an added partition [19]. Analyses of privacy hierarchy through visibility-graph methods report the same structure across houses of markedly different plan geometry [17].

The point of describing this configuration in syntactic terms is that it is measurable. Depth from the street, visual permeability and the position of the filtering sequence are quantities, not sentiments, and they can therefore be asked of a generated plan. Section 6 argues that current conditioning mechanisms cannot ask them.

4. CORPUS AUDIT

Table 1 records the documented properties of the four corpora that underpin the majority of published layout-generation research, and Figure 1 shows their combined scale against their combined climatic coverage.

Table 1. *Documented properties of the four principal public floor-plan corpora*

Corpus	Origin	Scale	Köppen class of origin	Dominant typology	Unroofed area labelled as
RPLAN [1]	China / East Asia	80,788 plans	predominantly Cfa / Dwa	Apartment, living-room hub	Balcony
LIFULL HOME'S subset [2], [22]	Japan	117,587 plans	predominantly Cfa / Dfa	Compact apartment	Balcony
Modified Swiss Dwellings [7], [23]	Switzerland	5,372 plans, 18,900 apartments	predominantly Cfb / Dfb	Multi-unit residential	Balcony, background
CubiCasa5K [6]	Finland	5,000 plans	predominantly Dfb / Dfc	Apartment and detached house	Outdoor, background

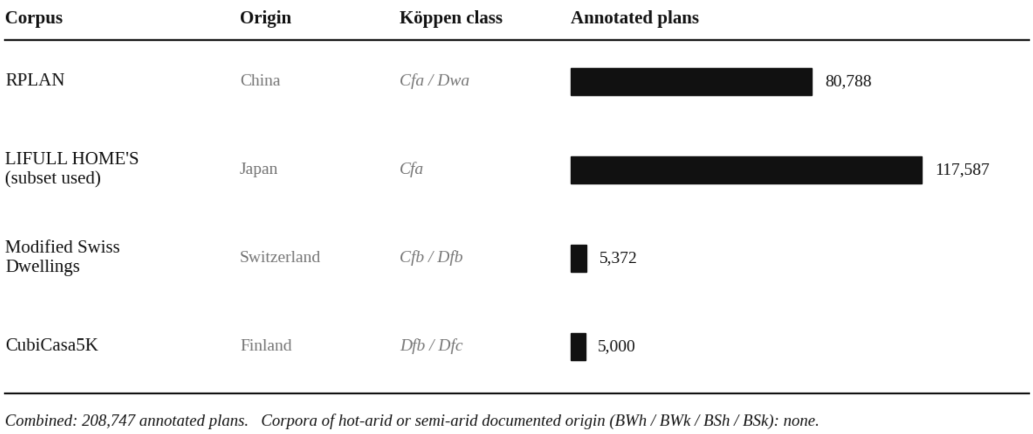

Fig. 1. *Provenance and climatic coverage of the principal public floor-plan corpora*

Three observations follow directly from the record. The first is climatic: the aggregate exceeds two hundred thousand annotated plans and documents no hot-arid or semi-arid origin. The arid and semi-arid belt includes many of the world's fastest-urbanising cities, and it contributes nothing to the reference distribution against which plan realism is currently measured.

The second is typological. Each corpus is dominated by extroverted dwellings in which an interior space — a living room in the Asian sets, a hallway or corridor in the Nordic and Swiss ones — acts as the circulation hub. This is not incidental to the models built on them: the seminal boundary-conditioned method explicitly positions the living room first and generates the remaining rooms iteratively from it, a design choice justified by the observation that a living room is present in almost all plans in the corpus [1]. A structural regularity of Chinese apartment stock is thus embedded in the generative procedure itself, not merely in the weights.

The third is schematic, and it is worth stating with precision because it is the claim most open to challenge. The corpora are not uniformly coarse: CubiCasa5K annotates over eighty object categories spanning rooms, fixtures and structural elements [6]. The claim here is narrower and survives that fact. In none of the four is an unroofed interior court a distinct spatial class. The published class vocabulary of Modified Swiss Dwellings is bedroom, living room, kitchen, dining, corridor, stairs, storeroom, bathroom, balcony, structure and background [23]; RPLAN, after the standard merging applied in downstream work, exposes eight room types of which the only exterior category is balcony [1], [5]; the graph-conditioned line built on the Japanese subset operates on a comparable ten-type vocabulary [2]. Wherever unroofed area occurs it is absorbed into balcony, outdoor or background.

The consequence is not that these corpora contain no open space. It is that open space appears in them only as a residual attached to a room, never as a spatial element that other rooms are organised around. A courtyard is therefore not an entity the resulting model can be asked to produce, refuse or evaluate: it is outside the ontology, and a reader can verify this from the published class lists rather than taking it on assertion.

4.1 Recent Corpora and Whether They Change the Finding

The audit above is open to a dating objection: the four corpora are those on which the published methods were built, and newer sets may already have corrected the imbalance. Two recent releases test the objection directly, and neither displaces the finding.

The first is a multimodal dataset of nearly twenty thousand floorplan images with associated metadata, curated from open internet sources expressly because prior work was confined to one building category and one country [24]. Its geographical spread is genuine and its motivation is close to the one argued here. It is nevertheless not a counter-example, for two reasons. It is a corpus for floorplan understanding — building-type classification, semantic *parsing, retrieval — rather than a room-typed layout corpus on which conditional generators are trained, and its heterogeneity is of building type and drawing format, spanning castles, temples and civic structures, rather than of residential typology within a controlled schema. Diversity that is not organised as a labelled residential ontology cannot be conditioned on, which is precisely the argument of Section 6.*

The second is a vector-graph corpus of seventeen thousand residential plans, released with room-connectivity graphs and explicitly positioned as a remedy to the limitations of the earlier sets [25]. It is directly comparable in kind to the corpora audited above, and it reproduces both properties at issue. Its functional categories are the familiar interior rooms together with balconies, so unroofed area again enters only as a residual attached to a dwelling rather than as an organising element; and its plans are drawn from public real-estate listings without a published geographical breakdown, so its climatic provenance cannot be established from the record at all. The second point is worth stating plainly: a corpus that does not document where its plans come from cannot be audited for climatic coverage, and undocumented provenance is not evidence of breadth.

The pattern across both is the same one the rest of this paper analyses. Collection has broadened; the schema has not. This is what the following sections predict, and it is why the argument is addressed to the encoding rather than to the sample.

5. FIVE REQUIREMENTS AND THEIR REPRESENTABILITY

From Section 3 we extract five requirements that a plan must satisfy for the hot-arid courtyard typology to function. Table 2 assesses each against the encodings in current use — pixel or raster room maps, room-adjacency graphs, and vectorised polygon sets — and distinguishes failures of data from failures of representation.

***Table 2.** Requirements of hot-arid courtyard housing against current plan encodings*

No.	***Requirement***	***What it demands of the plan***	***Expressible in current encodings?***	***Nature of the gap***
R1	*Void-centred organisation*	*An unroofed void is the primary organising element that all major rooms face*	*No: the unroofed court is not a distinct class in any corpus vocabulary*	*Representation*
R2	*Proportional geometry of the void*	*Height-to-width and length-to-width ratios within a performing range*	*Partly: geometry is present in vector encodings but the third dimension is not*	*Representation*
R3	*Seasonal migration*	*Opposing room groups occupied in different seasons within one envelope*	*No: plan encodings are static and carry no temporal use dimension*	*Representation*
R4	*Graded privacy sequence*	*Filtered depth from street to inner domain, with controlled sightlines*	*Partly: depth is a graph property but adjacency conditioning does not constrain it*	*Representation and evaluation*
R5	*Massive, low-aperture envelope to the street*	*Thermal mass and minimal external openings on exposed faces*	*Yes, in principle: wall thickness and openings are encodable*	*Data*

Only R5 is a data problem in the ordinary sense. The remaining four are not solved by adding examples, because the thing to be added has no slot to occupy. This distinction is the paper's central claim and Section 6 sets out its mechanism.

6. WHY ENLARGING THE CORPUS IS NOT SUFFICIENT

6.1 The Void Is Not a Room

Graph-conditioned generation takes as input a set of typed nodes and their adjacencies and produces room geometry consistent with that graph [2], [3]. The question is therefore whether the organising element of the courtyard house can be stated in that vocabulary. In the corpus typology the question does not arise, because the organising element is the living room, which is already a typed node. In the courtyard typology the organising element is an unroofed void, defined by three properties: it is not enclosed, it is not conditioned, and it is faced rather than occupied continuously.

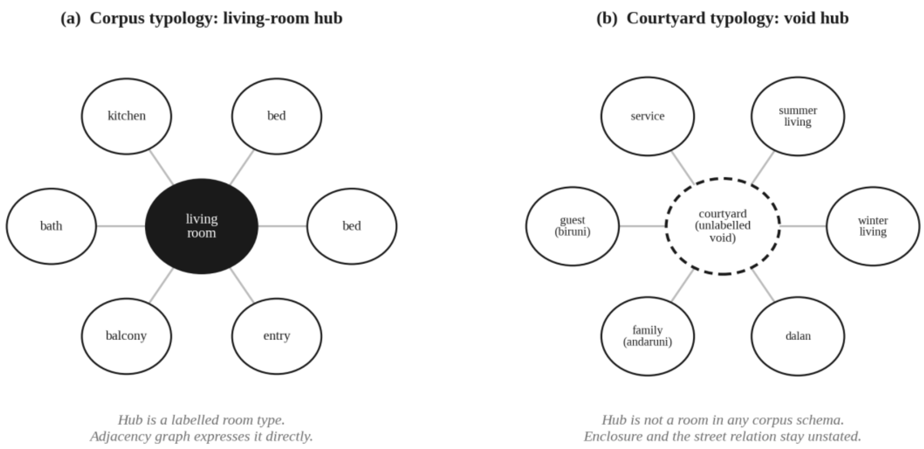

Fig. 2. *Hub structures compared: a labelled room type against an unlabelled void*

Figure 2 makes the asymmetry visible. One could of course add a node type called courtyard and retrain, and this is the obvious response. It is worth granting immediately what such a step would achieve: a star of adjacencies between a courtyard node and every major room is perfectly expressible in an adjacency graph, so the positive part of the requirement is not the difficulty. Two other parts are.

The first is negative. What distinguishes the introverted plan is not only that rooms face the court but that they do not face the street. In these formulations the exterior is not a node at all; the graph describes relations among rooms, and the boundary enters only as a raster mask or a polygon [1], [7]. A constraint of the form 'no habitable room opens to the public face' therefore has nowhere to be written, and a generator satisfying every declared adjacency may still open the plan outward.

The second is attributive. Adjacency carries no information about enclosure. A node labelled courtyard, joined to all its neighbours, is satisfied equally well by a roofed central hall, which is precisely the configuration the corpus statistics favour. Nothing in the objective distinguishes an unroofed void from an interior circulation space, and the difference between them is the entire climatic argument. The result is a plan in which the courtyard has become a light well beside the circulation rather than the element the house is organised around.

A third gap follows from the same source and is easier to state. The performance of the courtyard depends on the ratio of its height to its width, since shading and night-sky exposure are governed by the section rather than the plan [11], [14]. A floor plan, in any of these encodings, has no height. The dimension in which the typology does its climatic work is not merely poorly sampled; it is not present.

6.2 Depth Is Not Adjacency

The privacy gradient described in Section 3.2 is a statement about topological depth from the exterior. Depth of this kind is read from a justified graph, in which every space is placed at its shortest permeability distance from a chosen root [18]. Figure 3 contrasts the two configurations: in the corpus apartment the most private room typically sits at depth three from the street, while in the courtyard house the filtering sequence places the inner domain at depth five or more, and the additional steps are precisely the filtering elements.

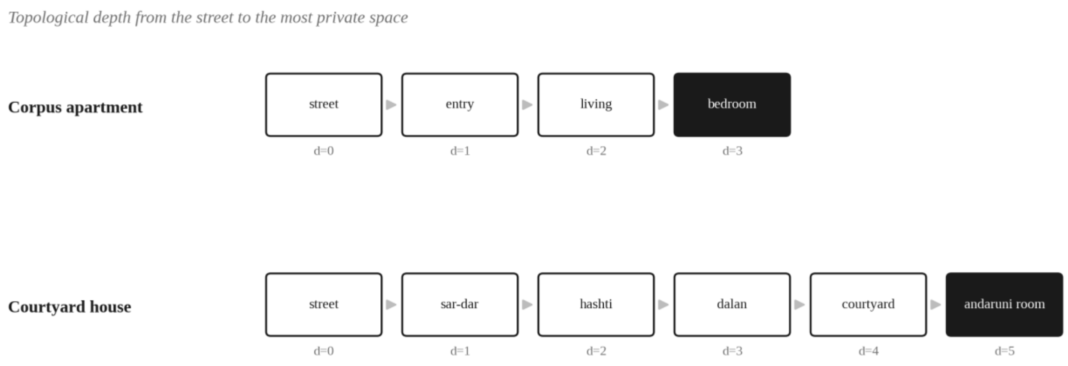

Fig. 3. *Topological depth from street to the most private space in the two typologies*

An adjacency graph cannot express this requirement, because depth is not a pairwise relation. Two plans with identical adjacency sets can have different depth profiles depending on where the entrance is placed, and in the corpora under discussion the entrance and the street are not nodes at all, so the root from which depth would be measured is itself unavailable. A model conditioned on adjacency is therefore free to satisfy its constraints while flattening the gradient — which is exactly the transformation that space-syntax comparisons record between traditional and modern housing in the same cities [15]. The generated plan is not wrong by the model's criteria; the criterion is missing.

6.3 Time Is Not in the Plan

Seasonal migration between winter and summer quarters is a defining behaviour of the typology and one of the principal mechanisms by which it achieves comfort without mechanical systems [13]. It is also entirely absent from every encoding under discussion. A plan in these corpora is a static labelled partition; there is no field in which to record that two rooms serve the same function at different times of year, and no loss term that could reward a layout for supporting the alternation. A generative model can no more learn seasonal use from these representations than it could learn acoustics from a floor plan. This is the clearest case in which the limitation belongs to the representation rather than to the sample.

6.4 The Evaluation Problem

A further problem compounds the three above. The metrics established for this task — realism against the corpus distribution, diversity, and compatibility with the input graph [2] — are all defined relative to the training corpus. A plan that departs from the corpus typology is therefore penalised as unrealistic by construction. Applied across regions, the standard evaluation protocol does not merely fail to detect the substitution of one typology for another; it actively rewards it. Any attempt to extend generative layout models to hot-arid contexts that does not first replace the evaluation criteria will report success while producing the failure it was meant to avoid.

6.5 Do Newer Architectures Escape the Problem?

An obvious objection is that the analysis targets graph-conditioned generative adversarial networks, which are no longer the frontier, and that language-conditioned or diffusion-based methods dissolve the difficulty. They do not, for three separate reasons.

Language conditioning changes the interface, not the ontology. A brief that asks for a house organised around a central courtyard is parsed into the same room vocabulary the corpus supplies, and the principal language-guided dataset for this task is constructed directly on top of an existing corpus, inheriting both its typology and its label schema [5]. The model becomes easier to instruct and no more capable of representing what the instruction names.

Diffusion methods trained on rendered plan images inherit the corpora more completely still, since a raster plan carries no explicit room ontology at all and the model's only access to typology is the statistical regularity of the *images it saw. Where such models are evaluated, the benchmark remains the same set of corpora, so improvements are measured as increasing conformity to the distribution whose locality is the problem.*

Most fundamentally, none of these advances alters the three representational gaps identified above. A void that organises the plan is still not a typed entity; depth from the street is still not a pairwise relation; and seasonal alternation is still absent from a static partition, whatever generates it. Capability has improved considerably along the axes the representation supports, which is precisely why the axes it does not support are becoming harder to notice.

7. THE CORPUS–CONTEXT DISTANCE PROTOCOL

The preceding analysis yields a practical instrument. If a generative layout model carries the assumptions of its corpus, and if those assumptions are invisible at the interface, then the minimum professional obligation is to state them before use. The Corpus–Context Distance protocol, shown in Figure 4, structures that statement along five dimensions.

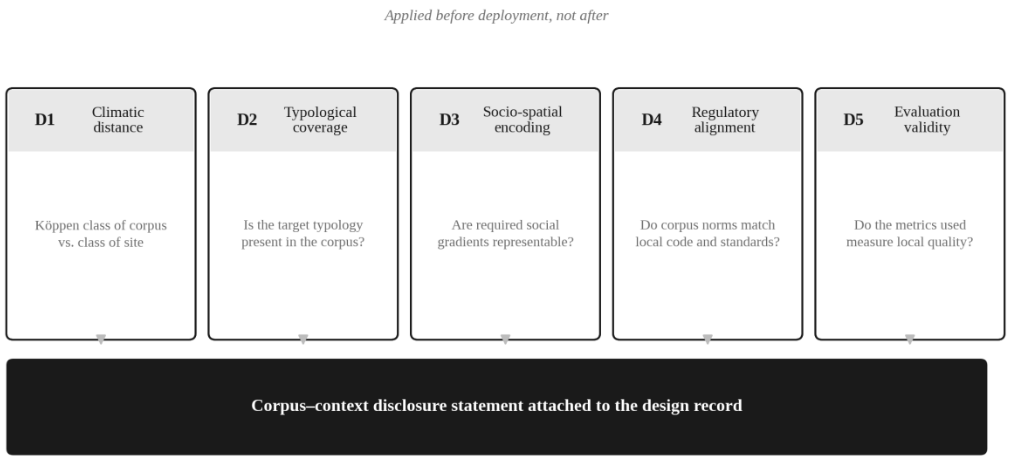

Fig. 4. *The five dimensions of the Corpus–Context Distance protocol*

Table 3. *Diagnostic questions and disclosure content by dimension*

Dim.	*Name*	*Diagnostic question*	*Content of the disclosure*
D1	*Climatic distance*	*Does the Köppen class of the corpus origin match that of the site?*	*Corpus origin classes, site class, and whether passive strategies differ in kind*
D2	*Typological coverage*	*Is the target typology present among the corpus categories?*	*Whether the target typology can be named in the model's schema at all*
D3	*Socio-spatial encoding*	*Can the required social gradients be stated as constraints?*	*Which gradients were specified, which were checked manually after generation*
D4	*Regulatory alignment*	*Do corpus norms conflict with local code, standards or tenure?*	*Identified conflicts and how each was resolved in the delivered design*
D5	*Evaluation validity*	*Do the metrics applied measure quality as locally defined?*	*Metrics used, their corpus dependence, and the substitute criteria applied*

Three conventions govern use. The protocol is applied before deployment, since a disclosure written after a design is complete is a rationalisation rather than a control. It is applied per project rather than per tool, because the same model has different distances from different sites. And a large distance is not a prohibition: it is a statement of the verification burden the practice has accepted, which is the form in which such judgements can be reviewed by a client, an insurer or a jury.

7.1 A Worked Example

The example below is a worked demonstration of the instrument, not a project report: no design was produced and no practice is described. It is constructed so that every input to the protocol is drawn from the published record and can be checked against it — the corpus properties from Table 1, the site classification from the Köppen system, and the typological requirements from Table 2. Consider a practice designing courtyard-based housing in Yazd, and adopting a graph-conditioned layout generator trained on RPLAN. D1 records a distance of maximum magnitude, from the predominantly Cfa and Dwa origin of the corpus to the BWk class of the site, with the consequence that the passive strategies implicit in the corpus differ in kind from those the site requires. D2 records that the target typology cannot be named: no courtyard category exists, so the practice must generate room groups and compose the void manually. D3 records that the privacy sequence was not expressible as a constraint and was instead checked by justified-graph analysis after generation, with candidates failing a minimum-depth criterion rejected. D4 records a conflict between corpus-typical window-to-wall proportions on the street face and local expectations of visual seclusion. D5 records that corpus realism was not used as an acceptance criterion and was replaced by a panel review against regional precedent.

What the practice can now say is not that it used AI, but that it used a model whose competence was documented, whose blind spots were identified in advance, and whose outputs were accepted on criteria appropriate to the place. That is a defensible professional position, and it is available today without any change to the underlying models. What the example establishes is the form of the disclosure and the fact that each of its five entries can be completed from information already in the public record; whether completing it changes the timing at which typological conflicts are detected is an empirical question, and P4 in Section 9 states the comparison that would answer it.

8. DISCUSSION

8.1 What an Adequate Corpus Would Require

The analysis implies a specification rather than a request for more data. A corpus adequate to hot-arid courtyard housing would need, at minimum, four properties absent from current sets: a room-label schema in which unroofed voids are first-class typed entities with recorded geometry; a third dimension sufficient to express the enclosure ratios on which courtyard performance depends; an annotation of seasonal or diurnal use where it exists in the documented record; and a topological annotation of the entry sequence, so that depth and filtering can be conditioned rather than inspected afterwards. The first and fourth are achievable from existing documentation of historical housing. The second requires measured heights or sections, which exist for surveyed heritage stock but not for the real-estate material from which current corpora are drawn. The third is the most demanding, since it requires ethnographic or archival evidence of use rather than geometric survey alone.

8.2 Beyond the Courtyard

The courtyard is a demonstration case, chosen because it is well documented and because its failure modes are sharp. The argument generalises to any typology whose organising principle is not a labelled interior room: the compound and the shaded veranda in humid tropical housing, the multi-generational compound, the elevated flood-adaptive dwelling of flood-prone deltas. In each case the question posed here is the right one to ask first — not whether the corpus contains enough examples, but whether the encoding has a place to put them.

8.3 A Note on Nostalgia

An objection deserves direct answer. It may be said that contemporary housing in hot-arid cities is largely apartment stock, that the courtyard house is a heritage form, and that a model trained on apartments is therefore *well matched to what is actually built. The observation is accurate and the inference does not follow. Contemporary housing in these regions is widely documented as having become dependent on electricity and gas for thermal comfort where the earlier stock was not [13], and space-syntax comparison of traditional and modern houses in the same city records the flattening of the access hierarchy that the earlier typology sustained [15]. A generative model trained on such stock does not merely reflect current practice; it entrenches it at the moment when climatic pressure makes revisiting the earlier logic most urgent. The argument here is not that new housing should imitate historical houses. It is that a design instrument which cannot represent the alternative cannot participate in the choice.*

8.4 Implications for Practice and Education

For practice, the immediate implication is procedural: the corpus of a generative tool is a project fact and belongs in the project record, in the same category as the version of a structural code. For education, the implication is that students working with generative tools should be examined on the interrogation of provenance — what corpus, what typology, what excluded — as a competence distinct from producing outputs. A student who can generate fifty plans and cannot say which typology the generator assumes has learned the less important half of the skill.

8.5 Limitations of This Study

Three limitations bound the claims. The audit covers publicly documented corpora only, and it rests on what those publications record rather than on independent inspection of every plan: where a corpus does not report the geographical distribution of its sources, the finding is that hot-arid provenance is undocumented, not that it is demonstrably absent. Vendors of commercial design software may likewise hold proprietary sets with different provenance, and the argument would need re-examination if such sets were disclosed. The typological analysis is grounded in the Iranian courtyard literature, which is unusually well documented, and the extension to other hot-arid traditions in Section 8.2 is argued rather than demonstrated. And the study infers model behaviour from documented properties of corpora and encodings rather than observing it; that inference is the substance of the propositions in Section 9, and if those propositions fail, the diagnosis offered here fails with them.

9. TESTABLE PROPOSITIONS

The study is documentary, and its value depends on whether its predictions hold. Table 4 states them in refutable form.

Table 4. *Propositions arising from the analysis and corresponding study designs*

No.	***Proposition***	***Study design***	***Disconfirming result***
P1	*Adding a courtyard node type to a graph-conditioned generator does not produce void-centred configurations at rates comparable to the historical typology*	*Retraining with augmented schema; configurational scoring of outputs against a documented sample*	*Augmented model matches historical configuration rates*
P2	*Plans generated under adjacency conditioning exhibit systematically lower street-to-private depth than the regional typology requires*	*Justified-graph analysis of generated versus surveyed plans*	*Depth profiles are statistically indistinguishable*
P3	*Corpus-based realism metrics penalise valid regional layouts*	*Expert panel rating of regionally valid plans, compared with the models' realism scores*	*Realism scores correlate with expert judgement*
P4	*Practices applying the Corpus–Context Distance protocol identify typological conflicts that would otherwise surface only after design approval*	*Comparative study of projects with and without the protocol*	*No difference in the timing of conflict detection*

10. CONCLUSIONS

Generative layout models are presented as instruments for architectural design in general. The audit presented here shows that they are instruments for a particular kind of dwelling in a particular band of climates, and that the particularity is invisible in use because the models fail gracefully: they return plausible plans everywhere.

The paper's argument can be stated compactly:

- *Over two hundred thousand plans in the principal public corpora, and none documenting a hot-arid or semi-arid origin; in none of them is an unroofed interior court a distinct spatial class, and the two most recent releases do not change this.*
- *Of five requirements of the courtyard typology, only one is a shortage of data; three fail in the representation itself, and a fourth fails in both representation and evaluation.*
- *Corpus-referenced realism metrics penalise regionally valid departures, so the standard protocol will report success while typological substitution proceeds.*
- *The Corpus–Context Distance protocol converts these findings into a pre-deployment disclosure that practices can adopt now, without waiting for better models.*

The wider point concerns what a design tool is. A generative model does not offer neutral capability; it offers a compressed account of how people have lived somewhere, and it offers that account wherever it is installed. When the account is unstated, an architect in a hot-arid city may adopt the spatial assumptions of a temperate apartment without ever deciding to. Making the account explicit is not a constraint on the technology. It is the condition under which the technology can be used architecturally at all.

REFERENCES

[1] Wu, W., Fu, X.-M., Tang, R., Wang, Y., Qi, Y.-H., Liu, L. (2019). Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics, vol. 38, no. 6, p. 1-12.

[2] Nauata, N., Chang, K.-H., Cheng, C.-Y., Mori, G., Furukawa, Y. (2020). House-GAN: relational generative adversarial networks for graph-constrained house layout generation. Computer Vision - ECCV 2020, Springer, p. 162-177.

[3] Nauata, N., Hosseini, S., Chang, K.-H., Chu, H., Cheng, C.-Y., Furukawa, Y. (2021). House-GAN++: generative adversarial layout refinement network towards intelligent computational agent for professional architects. arXiv preprint arXiv:2103.02574.

[4] Hu, R., Huang, Z., Tang, Y., van Kaick, O., Zhang, H., Huang, H. (2020). Graph2Plan: learning floorplan generation from layout graphs. ACM Transactions on Graphics, vol. 39, no. 4, 118.

[5] Leng, S., Zhou, Y., Dupty, M.H., Lee, W.S., Joyce, S.C., Lu, W. (2023). Tell2Design: a dataset for language-guided floor plan generation. arXiv preprint arXiv:2311.15941.

[6] Kalervo, A., Ylioinas, J., Häikiö, M., Karhu, A., Kannala, J. (2019). CubiCasa5K: a dataset and an improved multi-task model for floorplan image analysis. Scandinavian Conference on Image Analysis, Springer, p. 28-40.

[7] van Engelenburg, C., Khademi, S., Mostafavi, F., Standfest, M., Franzen, M. (2023). Modified Swiss Dwellings: a machine-learning-ready dataset for floor plan auto-completion at scale. 4TU.ResearchData.

[8] Srinivasan, R., Uchino, K. (2021). Biases in generative art: a causal look from the lens of art history. Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, p. 41-51.

[9] Mbalaka, B. (2023). Epistemically violent biases in artificial intelligence design: the case of DALL-E 2 and Starry AI. Digital Transformation and Society, vol. 2, no. 4, p. 376-402.

[10] El Moussaoui, M. (2025). Architectural practice process and artificial intelligence - an evolving practice. Open Engineering, vol. 15, no. 1, 20240098.

[11] Soflaei, F., Shokouhian, M., Mofidi Shemirani, S.M. (2016). Investigation of Iranian traditional courtyard as passive cooling strategy (a field study on BS climate). International Journal of Sustainable Built Environment, vol. 5, no. 1, p. 99-113.

[12] Givoni, B. (1976). Man, Climate and Architecture, 2nd ed. London: Applied Science Publishers.

[13] Izadpanahi, P., Mahmoudi Farahani, L., Nikpey, R. (2021). Lessons from sustainable and vernacular passive cooling strategies used in traditional Iranian houses. Journal of Sustainability Research, vol. 3, no. 3, e210014.

[14] Soflaei, F., Shokouhian, M., Soflaei, A. (2017). Traditional courtyard houses as a model for sustainable design: a case study on BWhs mesoclimate of Iran. Frontiers of Architectural Research, vol. 6, no. 3, p. 329-345. doi:10.1016/j.foar.2017.04.004.

[15] Alitajer, S., Molavi Nojoumi, G. (2016). Privacy at home: analysis of behavioral patterns in the spatial configuration of traditional and modern houses in the city of Hamedan based on the notion of space syntax. Frontiers of Architectural Research, vol. 5, no. 3, p. 341-352.

[16] Edwards, B., Sibley, M., Hakmi, M., Land, P. (2005). Courtyard Housing: Past, Present and Future. London: Taylor & Francis.

[17] Taheri Sarmad, F., Karami, A. (2024). Analysis of privacy hierarchies in traditional houses of Iran based on the plan structure through space syntax approach. Armanshahr Architecture & Urban Development, vol. 17, no. 47, p. 17-31.

[18] Hillier, B., Hanson, J. (1984). The Social Logic of Space. Cambridge: Cambridge University Press.

[19] Pirnia, M.K. (2004). Stylistics of Iranian Architecture. Tehran: Sorush Danesh. (In Persian).

[20] Jang, S., Roh, H., Lee, G. (2025). Generative AI in architectural design: application, data, and evaluation methods. Automation in Construction, vol. 174, 106174.

[21] Bender, E.M., Gebru, T., McMillan-Major, A., Shmitchell, S. (2021). On the dangers of stochastic parrots: can language models be too big? Proceedings of the ACM Conference on Fairness, Accountability, and Transparency, p. 610-623.

[22] Liu, C., Wu, J., Kohli, P., Furukawa, Y. (2017). Raster-to-vector: revisiting floorplan transformation. Proceedings of the IEEE International Conference on Computer Vision, p. 2195-2203.

[23] van Engelenburg, C., Mostafavi, F., Kuhn, E., Jeon, Y., Franzen, M., Standfest, M., van Gemert, J., Khademi, S. (2024). MSD: a benchmark dataset for floor plan generation of building complexes. Computer Vision - ECCV 2024, Springer. arXiv:2407.10121.

[24] Ganon, K., Alper, M., Mikulinsky, R., Averbuch-Elor, H. (2025). WAFFLE: multimodal floorplan understanding in the wild. IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), p. 1488-1497.

[25] Abouagour, M., Garyfallidis, E. (2025). ResPlan: a large-scale vector-graph dataset of 17,000 residential floor plans. arXiv preprint arXiv:2508.14006.