Graph-Conditioned Diffusion to Support Early-Stage Housing Layout Design

From Knowledge to Competitiveness
Portugal-US Partnerships 2026 Summit

MIT Portugal Program

Aydan Aghabayli [1]

aydanaghabayli@edu.ulisboa.pt

Supervisor: José Nuno Beirão[1]

1. Faculdade de Arquitetura, Universidade de Lisboa (FA-ULisboa), Lisbon, Portugal; CIAUD

MIT Portugal Program Doctoral Grant PRT/BD/154270/2022
Sustainable Cities and Urban Systems / Advanced Computing, including High-Performance Computing and AI

01 Research context and objectives

Decisions on dwelling **layout and facade** taken at the **earliest design stage** significantly affect the housing project's **cost** and **environmental performance** (Fig. 1).

Yet these criteria are usually addressed at later stages of design. A review of computational early-design tools shows that layout generation is typically conditioned on geometric plausibility or user-specified adjacency, not on cost or environmental criteria.

Objective

Develop a generative model that produces dwelling floor-plan layouts conditioned on facade performance targets, giving architects performance-informed layout alternatives to compare and adapt at the earliest design stage of affordable and sustainable housing.

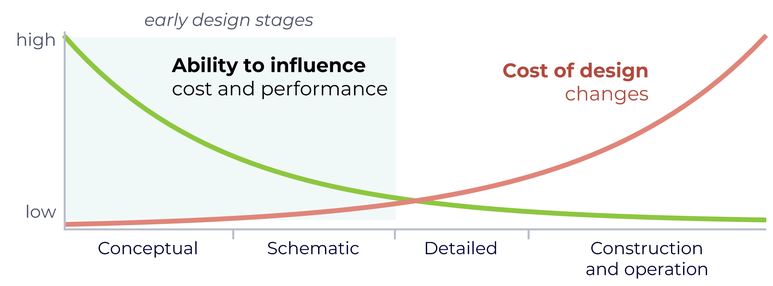


Fig. 1. Adapted from the MacLeamy curve (2004).

02 Housing corpus and data

The model learns from the **Modified Swiss Dwellings (MSD)** dataset [1]: real Swiss multi-unit housing with rooms, walls, openings and access relations annotated across thousands of floors.

5,372 floor plans

18,900+ apartments

18,634 dwelling graphs

From floor plate to dwelling

Floor plates hold ~32 rooms, many of them indistinguishable. Splitting them gives one graph per dwelling, 8.6 rooms on average, kept in the floor-plate frame so orientation stays true.

Fig. 2. Plan 9290: two apartments (blue, yellow) and shared stairs; lines link rooms by door.

Enriched room graph

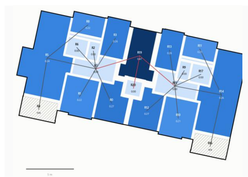
Each dwelling graph in the **Modified Swiss Dwellings (MSD)** carries node, edge and plan attributes (Fig.3).

Room (node)
type, area, ceiling height, window area and orientation, WWR (room), apartment

Door (edge)
width and type: entrance, passage, wet room, standard

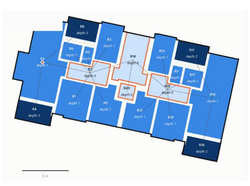
Plan (graph)
footprint area, perimeter, compactness

Compactness
heat loss, cost



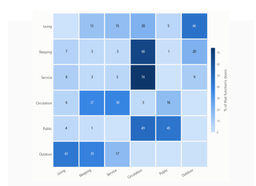
Window-to-wall ratio
per facade; daylight target

Depth to facade
daylight penetration



Depth from entrance
privacy, cross-ventilation

Balcony overhang
self-shading



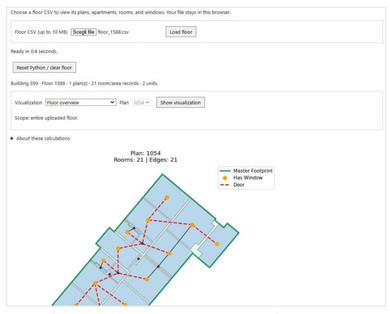
Room adjacency
circulation efficiency

Fig. 3. Six plan-level proxies, computed from geometry

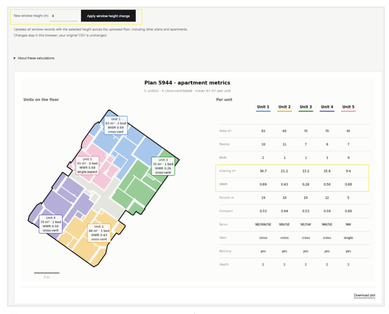
Data limitations. MSD records no U-values, HVAC or materials, and daylight was never simulated, so the performance target is glazing geometry (WWR). None of the major floor-plan datasets examined carries affordability-relevant attributes, which restricts conditioning to geometric and environmental targets.

03 Web application

A browser tool (Fig. 4) loads one floor as a CSV and renders **ten views** of these indicators. Python runs **in the browser**, so plans never leave the user's machine. Editing **window head height** updates WWR across the floor.



Loading a floor: room graph view



Per-apartment metrics

Fig. 4. Web application capabilities

04 Graph-conditioned diffusion

A **GATv2-based graph transformer** [2] encodes the room graph; a **denoising diffusion model** [3] generates rooms as **oriented bounding boxes**, conditioned on facade WWR targets.

Fig. 5. Reverse diffusion (left to right) refines each room's center, width, height and rotation from noise into a layout.

Training

- Losses: reconstruction, relative position, edge consistency, area, overlap
- Min-SNR timestep weighting [4]; plan-grouped splits prevent leakage
- Trained on the Deucalion supercomputer

05 Evaluation protocol

About **75% of pooled positional variance lies between dwellings**, so pooled scores mostly reward knowing where a dwelling sits. **Within-dwelling R^2** and **sample diversity** are the primary metrics, read against null baselines.

~75% between dwellings	**~25% within**

Only the within-dwelling share (teal) is scored.

06 Preliminary results

The model exceeds null baselines on within-dwelling room placement and on layout diversity (Fig. 6).

0.28
Within-dwelling R^2 room x-position

Model	**0.28**
Null	**0.07**

0.33
Layout diversity across samples

Model	**0.33**
Null	**0.04**

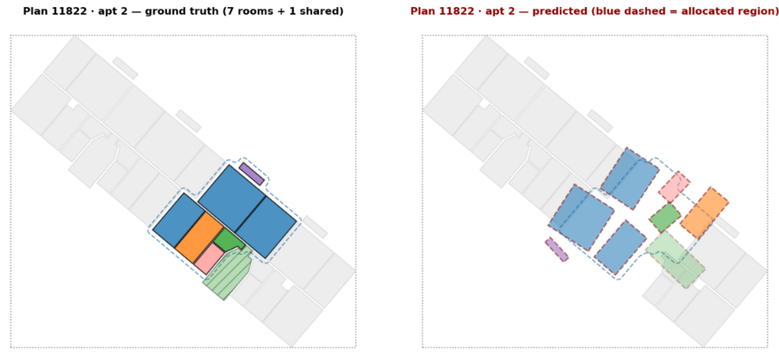


Fig. 6. Ground truth (left) and generated arrangement (right) for an unseen test dwelling; dashed outline = allocated unit region.

Next steps

- **WWR target sweep**: does the model causally use the conditioning signal?
- **Extend conditioning** beyond daylight to cost and embodied-carbon proxies.
- **Classifier-free guidance** for controllable generation.

07 Contribution and impact

- **Pipeline:** graph-conditioned diffusion linking layout generation to facade performance targets.
- **Protocol:** an evaluation that separates learned within-dwelling behavior from between-dwelling variance.
- **Data specification:** what data affordability-aware generative models need.

Lisbon. No machine-readable plan data exists. This work defines what such data must contain for these tools to support public housing design.

Acknowledgements Supported by the MIT Portugal Program through FCT doctoral grant PRT/BD/154270/2022. Computational resources were provided by the Portuguese Foundation for Science and Technology (FCT) through Advanced Computing Project 2026.03121.CPCA.A1 on the Deucalion supercomputer.

References [1] van Engelenburg et al. (2024) MSD: A benchmark dataset for floor plan generation of building complexes, ECCV. [2] Brody, Alon & Yahav (2022) How attentive are graph attention networks?, ICLR. [3] Ho, Jain & Abbeel (2020) Denoising diffusion probabilistic models, NeurIPS. [4] Hang et al. (2023) Efficient diffusion training via Min-SNR weighting strategy, ICCV.

Funded by:

MIT Portugal Program

Fundação para a Ciência e a Tecnologia

FACULDADE DE ARQUITETURA
LISBON SCHOOL OF ARCHITECTURE
UNIVERSIDADE DE LISBOA