

生成式人工智能在建筑设计领域的探索 ——以小库 AI 云为例

Exploration of Generative Artificial Intelligence in Architectural Design
A Case Study of LookX AI Cloud

[何宛余] HE Wanyu

[杨良崧] YONG Leong Shong Jackie (通讯作者)

作者单位

深圳小库科技有限公司 (深圳, 518000)

收稿日期

2023/06/30

国家重点研发计划项目 (2022YFC3801300)

DOI: 10.19819/j.cnki.ISSN0529-1399.202310006

摘要

以小库 AI 云为例, 探讨其如何整合生成式人工智能技术, 并介绍平台灵感创作、模型训练、共享社区和智能工作流。灵感创作使用大模型生成设计灵感图像, 模型训练则支持用户训练专属模型, 共享社区允许用户分享作品促进虚拟社区交流, 而智能工作流试图集成语言模型实现对话式交互设计。

关键词

智能化; 人工智能; 生成式人工智能; 小库 AI 云;
智能建筑设计

ABSTRACT

Using LookX AI Cloud as an example, this paper explores how it integrates Generative Artificial Intelligence technology, introducing inspiration generation, model training, community sharing, and AI workflows. The inspiration generation employs large models to generate design inspiration images, while model training supports users in training personalized models. The community sharing feature allows users to share their works, fostering communication within a virtual community. Lastly, the AI workflow attempts to integrate language models to enable conversational interactive design.

KEY WORDS

intelligentization; artificial intelligence; Generative Artificial Intelligence (GAI); LookX AI Cloud; intelligent architectural design

1 技术趋势

随着近期生成式人工智能 (Generative Artificial Intelligence, GAI) 相关底层技术的突破, 人类社会或许正在面临一个新的技术拐点。从人工智能 (生成式) 大模型发展历程来看, 随着硬件的发展, 训练模型可支撑的计算力在 21 世纪以来呈指数型增长, 这导致基于扩散模型的图像生成模型、可理解上下文的语言对话模型与多模态模型等在近几年实现^[1](图 1)。预计这将会很大程度上颠覆数个产业中涉及视觉、绘画、语言、口语等相关的工作场景。

当前也已涌现了数家深耕垂直领域的研发单位或科技公司, 试图将生成式人工智能技术应用于以往被认为离不开人类创意能力的生产过程。这些技术不再处于实验阶段, 而是已经可以在市场营销、计算机编程、音乐艺术、医疗医药、法律咨询等领域中找到可代表的商品化应用产品^[2]。在市场营销领域中, 市面上已有基于语言类的大模型, 如 ChatGPT 的软件应用工具可生成宣传文案; 在计算机编程领域中, 编程人员如今已不需要重新构思代码结构, 生成工具可以自主推荐潜在合适的代码编写样例; 在艺术创作领域中, 图像生成工具也已支持通过输入描述文字或参考图像生成各类图像作品, 其中包括逼真的照

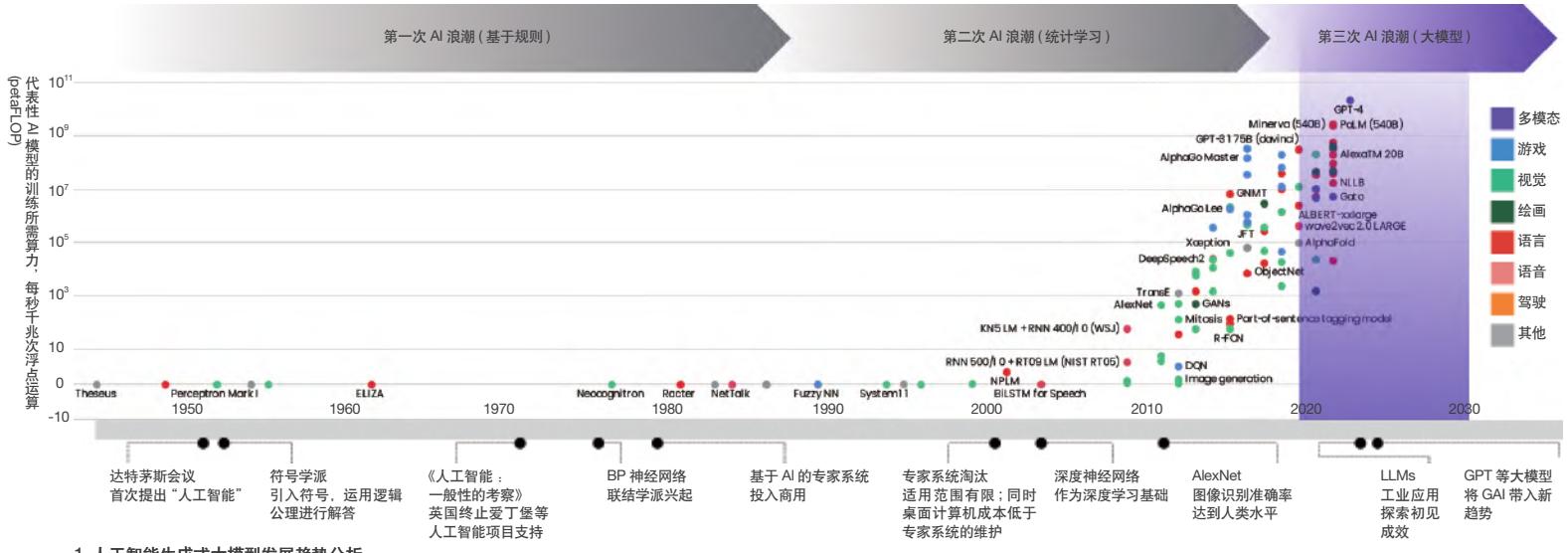
片和各类风格的插画。

在不久的未来, 生成式人工智能技术或许会更进一步向更高维度的数据格式拓展, 例如视频、三维几何模型等, 甚至是更多维的模型, 如建筑信息模型^[3]。而在应用领域方面, 该技术也被预测将可直接应用到需要更复杂的感知、创造和判断能力的工作场景中, 例如能实现文案和设计的最终稿、游戏平台的智能开发等^[3](表 1)。因此, “人工智能”这一原来被视为与创作、设计人员仍具有一定距离的词汇, 如今已经开始渗透到这些群体, 建筑设计也不例外。

2 建筑业趋势

社会和技术的演变都会催生新一轮的产业变革。建筑业自身也面临来自短期市场变化和长期新兴影响因素的冲击。市场变化对各行业的每一位参与者所产生的冲击较为直接, 但也由于其较容易受到突发、不可预期因素的影响, 从而难以深入讨论; 而对于具有长期效果的新兴因素则可展开探讨。

随着时代的演变, 建筑业目前面临 3 种趋势: 工业化、绿色化、数智化^[4]。工业化, 模块化制造、自动化制造、支持现场施工的自动



1 人工智能生成式大模型发展趋势分析

表 1 生成式大模型技术和相关潜在模态进展预测^[3]

时间	2020 年以前	2020 年	2022 年	2023 年	2025 年	2030 年
文字	垃圾邮件检测；翻译；基本问答 a	基本文案撰写；初稿 b	形式更长；草稿深化 b	垂直领域的细调取得更好的效果(科研论文等)c	终稿的质量高于人类平均水平 c	终稿的质量高于专业作者水平 c
编程	单行自动完成 a	多行生成 b	形式更长；准确率更好 b	适配更多计算机语言；适用更多垂直领域 c	文本生成(软件)产品 c	文本生成终版(软件)产品，质量高于全职的开发者水平 c
图像	—	—	艺术、标识、照片 b	初稿(产品设计、建筑设计等)b	终稿(产品设计、建筑设计等)c	终稿的质量高于专业艺术家、设计师、摄影师水平 c
视频、三维、其他	—	—	三维模型、视频生成的初步尝试 b	三维模型、视频的初稿 b	草稿深化 b	游戏的智能开发平台；个性化的视频游戏和电影 c

注：大模型可行程度(a 初步尝试；b 实现将至；c 迎来巅峰)

化技术将更好地为建筑业实现工业化，为达成“产品驱动（非项目制）”的建造方式成为可能。绿色化，在绿色低碳的概念下，不仅追求建筑在施工、运维层面的节能减碳，也呼吁绿色建材的应用。数智化，它与生成式人工智能的关系最深，包涵两层含义：1) 数字化，侧重对信息资源的形成与调用，基于信息技术所提供的支持和能力，让业务和技术产生交互，实现对信息的“获取和管理”；2) 智能化，侧重于对知识的获取并应用于工作过程，使机器具备感知、思维与判断、自适应学习、有效执行等功能，实现对知识的“提效和增值”。

因此，在建筑的数智化中，可实现更好的多方协作，更好地控制产业链，并支持数

据驱动的智能决策的制定。数智化也如同“工业 4.0”“智能制造”等概念，以往可能只有产业的主要参与者，或企业管理层特别关注，现如今这个认知已经悄然发生改变。当技术可以更多地为大众所获取和赋能时，掌握了该数字化和智能化能力的个体或组织，其创造价值的效率和效果将可以十倍、百倍于他者。反之，该个体或组织所创造价值的效率将远远落后于他者，所需的生产成本将会更高，并逐步被淘汰。人工智能技术的持续迭代下更加快了行业数智化的步伐。

3 建筑数字标准发展趋势

从全球的数字标准发展趋势中，我们或许能从中预测未来行业的发展方向。早在生

成式人工智能的横空出世前，国际标准化组织(ISO)于 2019 年成立了机器可读标准战略咨询小组(Strategic Advisory Group on Machine-readable Standards, SAG MRS)。他们试图对数字标准——“机器可用、可读、可迁移的标准”(Standards Machine Applicable, Readable and Transferable, SMART)进行初步定义。他们针对来自真实和虚拟世界的内容性质，并根据其与机器交互的不同程度，提出了 5 个等级^[5]：L1 文档化(纸质文本)；L2 电子化(开发数据格式)；L3 信息化(机器可读文档)；L4 数字化(机器可读内容)；L5 智能化(机器可交互内容)。其中，L5 智能化的数字标准包括了如下定义：1) 机器可定期自我学习的分析与校验，以改善内容的处理及访问；2) 机器可以查找和使用上下文中的元素，并对其执行操作；3) 机器可自动问答或预测并提供内容(图 2)。当将其投射到建筑领域可对应的内容范式时，可看到目前建筑业里还未有一种内容范式完全满足 L5 的定义，或者达到其目标效果，因此暂且将其称为“新一代的元数据”。理论上，L5 建筑数字标准不仅支持 L4 数字化所提到的“数”(非几何数据)和“模”(几何模型)等数据结构，还需引入“规”的概念，即涉及了规范、规则及规律。1) 规范：机器可应用标准、法规、规范；



2 ISO SMART 数字标准 (以建筑领域为例)

3 灵感创作：对设计草图生成灵感或渲染图像

2) 规则：机器可应用行业通识、企业业务的规则逻辑；3) 规律：机器可自主学习规律以改善内容及关系的处理，并自主问答及预测、提供内容。

L5 级智能化可能是建筑业未来与人工智能技术有机融合的载体指向之一，为此，需要能支持相应数字标准的内容范式、技术，甚至是相匹配的软件系统（工具）。下节将以小库 AI 云作为一个工具案例来展开介绍。

4 生成式工具的可能

小库 AI 云为深圳小库科技有限公司于 2023 年 5 月 7 日上线的一个设计云平台（以下简称平台），旨在回应当下的生成式人工智能技术发展和其在建筑设计领域的应用可能。平台提供了数个功能点以回应上述情况：“灵感创作”实现通过文生图、图生图的方式智能生成灵感图像；“模型训练”提供了低门槛的人工智能模型训练功能；“共享社区”则允许创作者在一种数字社区中分享作品。而“AI 工作流”则是试图回应当前对话式大型语言模型 (Large Language Models, LLMs) 的技术发展而提出的新设计流程可能。

4.1 灵感创作

鉴于图像生成模型的能力提升，生成更逼真的、能启发设计灵感的图像已成为可能。该平台中的“灵感创作”主要响应了设计前期阶段的如下痛点：

1) 项目和案例研究完成后确定了设计方

向或设计草图，但难以将设计关键词或草图转化为空间视觉想象；

2) 设计早期阶段制作的、写实度高的可视化图像虽可更好地表达设计概念，但为此需求进行渲染效果并不明显；

3) 难以根据某个指定的设计方案再拓展出不同幅度的设计变异可能性。

灵感创作允许设计师输入文字、草稿图像，甚至是成果参照图像来生成灵感或渲染图像。为了回应上述痛点 1) 和 2)，此功能针对不同场景的需求主要提供了两种模式：“轻松模式”和“探索模式”。轻松模式具有较低的使用门槛，适用于初学者，用户可选择软件中已预设好的生成场景，可上传多种类型图像作为用于生成构成参照的基础图像，并选择输入文字描述以生成灵感或渲染图像。探索模式则开放了更多的控制参数设置，允许用户从零开始，即不需上传任何参考图像，仅需输入文字描述，或选择数个标签智能组合描述句子，接着指定基础和风格的人工智能生成模型，并自行调节模型参数以生成最契合的灵感效果图。

以轻松模式为例，在实际的渲染作业中，用户以往需要对该设计方案搭建三维模型、设置渲染参数并进行后期图像处理，这个过程需要一定的时间成本。而使用此平台时，用户只需要上传某一张设计草图、三维素模截图、手绘素描、真实照片或线稿平面图来作为基础图像，接着通过一段文字描述，如

“写实渲染，独栋建筑，外观具有台阶式和绿化的公共空间或体块，现代风格”，或上传一张风格参考图像，点击生成按钮等待数分钟即可获得数张“灵感图像”。该生成的图像在画面构成上将会如实遵从基础图像，而在风格上将会参考文字描述或者风格参考图（图 3）。在细节上，虽然生成的图像无法完全媲美于那些通过常规的、对三维模型精度高的渲染技术手段所可达到的高精准度，但是经过此平台的智能“像素”演绎，大部分的画面构成、建筑材料和光影效果等亦可达到设计前期的沟通要求，甚至能支持下载不同放大尺寸倍数的高清图像。

在设计的发展过程中，基于某个设计方案进行拓展是一种常见的设计思路。针对痛点 3)，需要有别于传统地重新构思和调整设计模型，灵感创作还提供了另一种更便捷的方法，即用户只需要指定某张已生成图像，一键点击生成数张衍生图像以拓展思路。

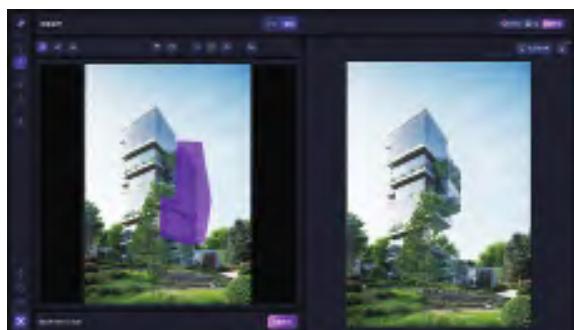
而对于任意一张已生成图像，此平台也支持局部编辑，即二次生成。以往用户需要重新局部调整三维模型，或切换到其他图形处理软件进行后期处理。此平台则允许用户通过多边形或涂抹的方式指定需要局部编辑的范围，接着仅需要输入文字描述，如“细化更多的立方体块”对该范围进行二次创作，改变建筑材料、风格、元素，甚至是形体（图 4）。

以上灵感图像的生成，建立在建筑垂直领域的底层 AI 大模型之上，包括了建筑大

模型、室内大模型和“泛建筑”大模型。这些模型之所以得以训练和应用，离不开千万级建筑结构化数据库——ArchiNet。该数据库具有中文语料库和标注，使得该大模型能提供高质量的建筑灵感生成。也基于此特点，该模型与市面常见的模型具有不同之处。后者由于为了适应更广泛的应用场景，往往容易陷入“过拟合”，即无法具体地理解某个行业术语，从而影响图像生成质量（图 5）。

表 2 模型平台使用特点对比

	平台名称	MJ 云平台	SD 本地部署	小库 AI 云平台
图像生成对比	使用门槛	硬件要求 低	高	低
	上手难度	低	较高	较低
	功能丰富性	内容质量 较高	良莠不齐	较高
	是否针对行业优化	否	可以，需要训练模型	是，小库自有的建筑与室内大模型
	支持中文提示	否	否	是，中英双语训练集，允许用户中文提示
模型训练对比	严格可控生成	否	可以依赖第三方插件	是，小库平台内置
	模型社区	无，官方只有少数模型	需要收集资源存储到本地	有，小库平台内置行业最大的建筑模型分享社区
	扩展性	支持行业语义 较少	是	是
	自定义风格训练支持	否	是，不过门槛极高	是，零门槛训练
	平台名称	某在线平台	SD 本地部署	小库 AI 云平台
模型训练对比	使用门槛	硬件要求 低	极高(16G 显存属于入门水平)	低
	上手难度	较低	高	较低
	功能丰富性	训练效果 较差	良莠不齐	较好，官方预设多个行业训练模板，多场景训练最优解
	参数自定义	不支持	极高	支持，提供两种模式，满足不同用户需求
	成果分享	支持	否	支持，可发布至模型集市，让所有人欣赏与使用您的成果
训练成果	训练成果版权	属于平台	用户自有	用户自有
	训练成果可下载	需额外操作	是	是



4 局部生成：一键智能局部修改图像

4.2 模型训练

1990 年代以前的设计工具是尺、规、纸、笔，而往后则有了电子化的计算机辅助设计工具。而在本世纪初后则有了建筑信息模型软件。随着参数化设计软件和可视化编程工具的出现，设计师可以开始创造自己的设计工具，特别是难以直接描绘的非线性形体，或者是需要重复操作的绘制和建模工作。

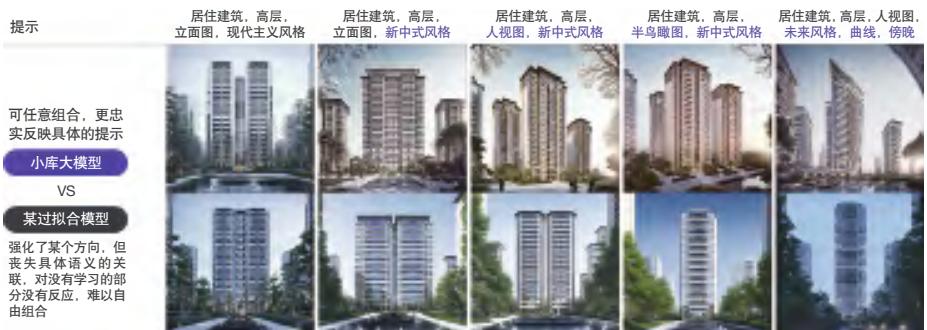
在生成式人工智能时代下，除了使用

现成的模型，更低门槛地创造工具成为可能，甚至是更“得心应手”。该平台除了提供官方标准的生成式模型，其中的“模型训练”功能也允许用户自行训练专属的生成式模型，作为后续创作的有力工具。这是凭借着上述 ArchiNet 数据库作为一种“泛建筑”的模型训练底层架构，同时该数据库也支持了中文标签，因此该平台的模型训练可更好地贴近国内的应用情况，而这正是国外常见的生成式模型所缺失的特点（表 2）。

为了满足不同的用户群体，该模型训练如同灵感生成，也分为自动模式和进阶模式（图 6）。如同灵感创作，自动模式具有较少的参数设置，而进阶模式则开放了更多模型训练的参数设置，其中包括训练步数、学习率、网络维度等。

以自动模式为例，在实际操作中用户需先整理 15~100 张图像（共同表达相似的元素，如某个具体的建筑设计手法、形体、材质、环境元素等）作为模型训练数据集上传到此平台。接着设置准确的训练集类型、训练方向等基本参数以便得到更精确的训练效果。经过大约 10min 的等待时间，用户即可获得专属的生成式模型。为了更好地表达该模型的实际效果，平台还会提出若干随机的文字描述，和不同等级的模型相似度强度作为两组生成输入条件进行自动化测试。该测试的生成图像，按照前述两组测试条件进行排列，形成一份直观的矩阵表格（图 7）。

当然这些训练成功的生成式模型可接着应用到前述的灵感创作中。它们能随时根据设计需要，随时被搜索和调用，实现重复的



5 小库大模型与过拟合模型效果比较



- 6 模型训练的自动和进阶模式
7 生成式模型自动化测试结果矩阵
8 灵感广场：显示平台可生成多类设计领域的图像，包括城市、建筑及景观
9 模型集市：显示平台具有多种设计领域的可选生成式模型
10 首批入驻该平台的模型训练团队：AARG (AI 建筑研究室)、樊川、XSarchitectural 和徐睿鹏工作室

设计逻辑和表达的复用。某种程度上，这似乎解放了创意层面的“重复劳动”。

4.3 共享社区

任何行业的设计成果都具有其价值，相应的知识产权保护也逐渐成为设计领域重要的关注要点。在此平台中，由用户自己创作的作品，即生成图像和人工智能模型的所有权将完全归属于用户，这一举措的目的之一是为了弱化用户在分享其作品后可能所导致的知识产权侵权隐患。基于此，任何用户可将生成的灵感图像和已训练的生成式模型个别分享到相应的线上公共社区，即“灵感广场”和“模型集市”。

在灵感广场中，所有平台用户可看到被允许公开展示的图像作品。在一种全开放权限的情况下，某位用户不仅可在“灵感广场”看到某件图像作品，还可一键生成相似“同款”图像（图 8），即直接复用生成该图像的模型和参数设置，甚至基于该模型的设置进

行调整以达到想要的生成效果。创作者可进一步设置权限，例如是否允许他人复用该图像生成的所使用的文字提示，或是复用该图像的生成式模型。

而在模型集市中，需求方也可以搜索、筛选、排序并直接使用已训练好并授权可复用的生成式模型（图 9）。在这里，具有编程能力的用户也可选择自行上传他们于本地或非此平台计算环境所训练的生成式模型文件，以供社群使用，提升“工具”的复用次数。由于在使用相同的模型和其参数设置下难以生成完全一样的图像，因此抄袭问题也难以存在，更多的是人机共创的成果被更多的人分享和“借用”。

社区的概念还能进一步拓展成市场，意味着该平台中的参与者扮演了需求方和供应方，在双方自愿的条件下，作品都可发生交易。例如需求方需要付费使用他人的人工智能模型，而该模型的创作者则可按调用次数

获得分成。这种交易机制的设计，在某种程度上也是为如今处于数字游民、大模型时代下的设计群体提供了另一种“业务思路”。由于模型训练的门槛降低，设计师不仅能提供设计服务，那些能自主训练出优秀模型且被多次调用的“炼丹师”也能通过此类新兴的模型训练工作获得收益（图 10）。

4.4 未来可能——智能工作流

前述三种功能主要可作用于设计前期阶段，尤其是设计灵感的构思和视觉具象化的作业场景中，这些功能的实现主要基于生成式图像大模型，而大型语言模型在设计领域中也有其可发挥的地方。后者通过学习大量的文本数据，可理解语言文本的含义（看懂问题），并生成如同人类自然语言的文本（回复问题）。由于此类模型使得机器可以更好地理解人类的提问，它成为多模态（multimodal）技术中最重要的一环。在多模态模型中，可通过输入多种格式，如语言、



11 規范助手界面



12 智能工作流设想

图像、视频等，接着输出多样数据类型。

ChatGPT 的横空出世无疑影响了整个世界，它向世人证明了大型语言模型某种程度上可以理解人类的自然语言，并做出逼真的反应。当然，此类数据驱动模型的局限也如同前述提到的市面上存在的图像生成模型，其“专业属性”依然难以完全适用于所有生产性工作场景。该模型主要使用互联网上的文字信息作为训练集，很多未上网的信息自然就未被纳入到其学习范畴。专业领域数据和知识的缺乏，结果是该模型固然看似具有广泛的知识，但还远未能达到某领域的专家水平。从人工智能发展历程可见，虽然人工智能从古至今或多或少也有应用于个人的娱乐场景，但是历代智能模型或学派的“生存”，是需要考虑该模型是否能产生价值？是否能成为当下不可或缺的生产力？抑或是能否提升现有的生产力？（图 1）

在垂直领域如建筑行业中已有一定规模的数字信息沉淀了下来，例如文字方面，其中典型代表有标准、规则、规范类的知识。例如一个学习了消防规范的模型被训练后，可提供更便捷的规范查询服务，小库 AI 云的“规范助手”即是源于这种尝试。用户无需手动查找规范文件及细项条文，仅通过与机器对话便能获得规范相关的结构化信息（图 11）。

对话式的交互具有其便利性，可直接表达人类的思维方式。受此类交互模式启发，将其集成到现有设计的流程可被设想为：在

上传住区设计项目任务书后，对话机器人可自主回答并推荐后续设计步骤，其中包括对该项目所在地推荐户型，生成可直接编辑的户型模型和建筑楼栋，推荐数个规划组合策略，并筛选满足合规性、目标和约束条件的住区规划方案。接着还允许机器对所选择的方案体量轮廓进行重新生成或人机互动编辑，最终得到满意的作品；在前述的“灵感创作”生成带有细节元素的渲染示意图，甚至再将此效果图转化为二维图纸或三维模型（图 12）。

5 结语

小库 AI 云的案例，相信只是生成式人工智能于建筑设计应用的一隅。以上案例探讨了此新技术目前更多可能应用的设计前期阶段，从灵感生成到自主推荐操作步骤。除了为建筑设计领域带来了令人兴奋的成果，该技术也可为设计师探索新的设计可能性，并有潜力彻底改变现有的建筑设计方式。

当然，面对任何新兴技术，在完全理解其风险前都必须谨慎对待，确保它的工作和输出成果仍然是可控的，尤其是认识并解决与其实施相关的道德和社会影响至关重要，例如在该平台的生成机制中就需要考虑对敏感词的处理。另外，模型算法的公正性也取决于它们所训练的数据，如果不仔细管理，生成结果有可能会产生偏见和不平等的情况发生。因此，在一个能公开分享图像和模型的线上社区中，软件开发者和用户都必须保

持警惕，确保其负责任和合乎道德地使用，同时考虑隐私、安全和社会影响问题。最后，生成式人工智能技术在建筑设计中的整合不应削弱人类创造力和直觉的作用。建筑是一门深刻的人文学科，涉及对文化、社会和情感背景的理解。其应该成为增强人类能力而不是取代人类能力的工具。人类专业知识、艺术视野和技术进步的融合将真正塑造建筑设计的未来。■

参考文献

- [1] Computation Used to Train Notable Artificial Intelligence Systems[EB/OL]. (2023-06-30)[2023-06-25]. <https://ourworldindata.org/grapher/artificial-intelligence-training-computation>.
- [2] BUHLER K. Generative AI Is Exploding[EB/OL]. (2023-04-11)[2023-06-25]. <https://www.forbes.com/sites/konstantinebuhler/2023/04/11/ai-50-2023-generative-ai-trends/>.
- [3] GRADY P, HUANG S. Generative AI: A Creative New World[EB/OL]. (2022-09-19)[2023-06-25]. <https://www.sequoiacap.com/article/generative-ai-a-creative-new-world/>.
- [4] RIBEIRINHO M J, MISCHKE J, STRUBE G, et al. The Next Normal in Construction: How Disruption Is Reshaping the World's Largest Ecosystem[R]. McKinsey, 2020.
- [5] 刘曦泽, 王益谊, 杜晓燕, 等. 标准数字化发展现状及趋势研究 [J]. 中国工程科学, 2021, 23(6): 147-154.

图表来源

图 1: 数据来自参考文献 [1]，笔者负责数据可视化
表 1: 参考文献 [3]
其余图表均为作者绘制