arXiv:2603.03074v1 [cs.HC] 3 Mar 2026

# 面向从业者的生成式人工智能设计：探索与创意实践相契合的交互方法

Xiaohan Peng Wendy E. Mackay Janin Koch LISN LISN UMR 9189 CRIStAL 巴黎萨克雷大学、法国国家科学研究中心、法国国家信息与自动化研究院 巴黎萨克雷大学、法国国家科学研究中心、法国国家信息与自动化研究院 里尔大学、法国国家信息与自动化研究院、法国国家科学研究中心、里尔中央理工学院 法国奥赛 法国奥赛 法国里尔 xiaohan.peng@inria.fr wendy.mackay@inria.fr janin.koch@inria.fr

## 摘要

设计是一种非线性、反思性的过程，从业者在其中运用视觉、语义及其他表达性材料来探索、迭代并完善创意。随着生成式人工智能（生成式人工智能）融入专业设计实践，传统的交互方法侧重于提示词或整图操控，可能导致人工智能输出与设计师的意图不一致，迫使视觉思维者转而进行口头推理或事后调整。我们介绍了*DesignPrompt*、*FusAIn*和*DesignTrace*中的三种交互方法，它们将控制权分配到意图、输入和过程各个方面，使设计师能够在交互的不同阶段引导人工智能对齐。我们进一步指出，对齐是一种动态协商过程，人工智能会根据设计师在工具性和启发性方面的需求以及创作阶段，采取主动或被动的角色。

## CCS 概念

• **以人为中心的计算** → **交互系统与工具**；•**应用计算** →艺术与人文。

## 关键词

创造力支持工具、人机交互、设计实践、生成式人工智能

**ACM 参考格式：**

Xiaohan Peng, Wendy E. Mackay, and Janin Koch。2026年。《面向从业者的生成式人工智能设计：探索与创意实践相契合的交互方式》。载于*2026*年 *CHI*计算机系统人因会议（*CHI '26*）双向人机对齐研讨会，*2026*年*4*月*13*日至*17*日，西班牙巴塞罗那。ACM，美国纽约州纽约市，4页。

## 1 引言

设计是一场“与材料的反思性对话” [18], ，从业者借助视觉、语义及其他表达媒介来引导创作过程。在专业实践中，创作很少是从概念到执行的线性路径；相反，它是一个交织着探索、反思与完善的过程[6]。设计师通过重用材料、精确掌控细节，以及借助草图和注释记录创意的演变，不断推进自己的工作 [4, 19]。

然而，随着生成式人工智能（生成式人工智能）在设计实践中日益普及，其交互范式往往无法充分使设计师的表达和意图与系统的响应相契合。主要以对话驱动的交互范式迫使以视觉为导向的设计师从视觉思维转向制定口头指令 [6, 13]。此外，设计词汇与生成式人工智能术语之间仍存在概念上的差距。例如，对设计师而言，“变体”是一种有意的迭代，用于探索设计空间 [11]；而对生成式人工智能系统而言，变体往往仅仅是扩散过程中不同随机噪声所产生的结果。设计师常常不得不接受一个起初看似合理但并不理想的“变体”，然后再进行手动调整 [7]。

现有视觉交互方法主要作用于完整图像输入，并可选择使用蒙版来指定局部细节控制。然而，这种交互范式预设了相对完整的视觉素材已经存在。它迫使设计师“修复”现有素材，而不是从零开始“创作”，使重点从创造性表达转向事后调整 [22]。此外，设计知识包含隐性抽象、层级结构和实际约束 [20] ，而这些内容很难通过执行文生图或图生图工作流来捕捉。核心问题不在于生成式人工智能提供了哪些模态，也不在于设计师如何沟通，而在于如何构建交互，以引出并捕捉设计表达，同时保留有意义的设计意图、过程和成果。

尽管一些系统通过参数化组件提供控制，但这类组件往往要求专业技术知识，且与创作决策之间缺乏直观对应关系 [3]。ComfyUI [1] 和 Stable Diffusion WebUI [2] 等更先进的界面提供了更高的灵活性，但代价是进一步增加了技术复杂性。设计迭代是一个有目的、非线性且依赖情境的过程，无法通过一维参数充分呈现。这凸显了对交互方法的需求：让设计师能够在自己偏好的抽象层级上引导人工智能对齐，同时不必放弃现有工作流。

在这篇立场论文中，我们介绍了生成式人工智能应用于专业设计实践的三种不同交互方法，展现了可供探索的交互多样性和设计空间。随后，我们探讨人工智能在人与人工智能协作设计流程不同阶段中所发挥的作用。

## 2 面向创意实践的交互方法

我们认为，设计师与生成式人工智能之间的对齐并非通过优化交互中的某个单一环节——无论是输入、输出，还是生成本身——即可实现的静态目标，而是

[1] https://www.comfy.org/ [2] https://github.com/AUTOMATIC1111/stable-diffusion-webui

本作品采用知识共享署名 4.0 国际许可协议授权。*CHI '26*，西班牙巴塞罗那
© 2026 版权所有，归作品所有者/作者所有。

在创意实践的完整历程中动态形成。当前的交互方法通常将对齐定位于单一环节，最常见的是提示词。然而，专业设计实践中的意图、行动与反思分布于不同的时间、空间和表征之中。我们提出了三种交互范式，将交互中的结构安排在不同位置：融入意图的多模态分解（*DesignPrompt*）、融入对视觉提示词的直接工具化操控（*FusAIn*），以及融入设计过程本身的时空记录（*DesignTrace*）。这些范式共同为设计师提供了多个着力点，使其能够引导生成式人工智能的生成能力，产出有意义的创意成果。

## 2.1 意图对齐：可配置的多模态支架

在近期的人机交互研究中，围绕设计相关维度来结构化并搭建生成式人工智能输入的支架，已成为一种能够更好地表达设计师意图的交互方法。此类输入方法将输入分解为有意义的单元，例如设计属性、视觉参考或语义维度，供设计师配置和重新组合 [2, 12, 21, 23]。

一个这样的例子是 DesignPrompt [15]，它让设计师将意图分解为离散且可操作的模态：图像、颜色和语义标签（图 1）。这些可组合的元素构成多模态脚手架，使设计师的视觉意图变得明确且可单独调整。该系统同时透明呈现多模态输入以及由此产生的转换后的提示，使设计师能够在将输入发送给生成式人工智能之前完善其输入。通过将意图拆分为交互式组件，设计师可以保持某些属性（例如配色方案）不变，同时沿其他维度探索变体，从而在其所设想的内容与系统所生成的内容之间实现更精确的对齐。

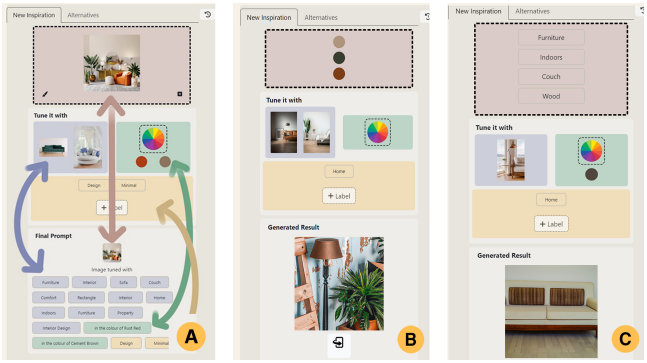


**图1：DesignPrompt通过多种输入模态支持意图分解：（A）图像修复、（B）颜色，以及（C）语义标签。多模态输入被映射为可编辑的文本提示词，从而支持交互式提示编辑与重新排序，以优化提供给生成式人工智能的文本输入。**

## 2.2 输入对齐：可直接操控的视觉提示

当设计师无法直接介入创作行为时，对齐往往会失效。近期研究探索了生成式人工智能系统，使从业者能够更主动、更直接地交互，从单次生成转向设计师可以迭代塑造视觉输出的界面。一些系统通过让设计师在不同粒度层级进行输入 [17]，并在高保真生成的同时持续绘制草图 [8, 9]来解决这一问题。

我们围绕 FusAIn [16] 开展的工作，通过将文本和图像提示词具象化为可加载特定视觉属性（如纹理、颜色或物体）的工具型“智能”画笔，并通过直接操控将其应用于画布上（图 2），拓展了这一领域。这一做法将输入从静态描述转变为富有表现力的空间操作：设计师不再通过描述期望结果来调整生成过程，而是通过笔触和放置亲手构建目标内容。这种方法将交互从生成与验证循环转变为持续、细致地对难以仅凭文本或图像表达的细节进行视觉引导。

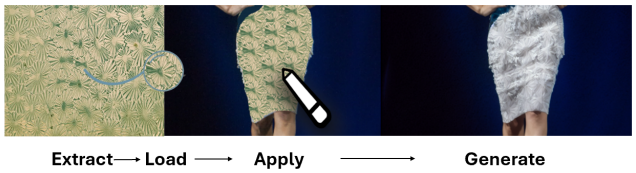


**图2：FusAIn 将基于图像的视觉提示具象化为个性化“画笔”，用以编码特定视觉属性。使用“纹理画笔”时，用户从源图像中提取纹理，将其载入画笔，再通过绘制将纹理应用到目标视觉内容上。随后，系统生成图像，在保留其他属性的同时，始终如一地应用该纹理。**

## 2.3 过程对齐：迭代式、可追踪的生成式工作流

从业者与生成式人工智能之间的创意对齐很少能在一次生成中实现；它是在随时间展开的非线性探索、比较与选择性完善循环中逐步形成的。然而，这种对齐容易受到累积性生成漂移的影响：与生成式人工智能反复迭代，往往会偏离设计师最初的意图 [14]。由于每次生成都可能对先前的输出进行“破坏性”编辑 [1]，，若未能记录这些编辑之间的转变，就无法追溯创意决策如何演变、各种方案之间有何关联，或为何选择了某个特定方向。

DesignTrace [14] 引入了一种可迭代、可追踪的节点式交互方法，用于捕捉并结构化与生成式人工智能协作的非线性设计过程轨迹（图 3）。通过将设计依据具象化为视觉和语义制品，并使这些轨迹在时间和空间上均可追踪，该系统帮助设计师在不同的探索方案和迭代中保持视觉一致性。系统通过结构持久性实现对齐：设计师可以回顾早期状态，在不丢失

上下文的情况下开辟新的创意分支，并观察各次迭代之间的联系，确保人工智能的演进轨迹始终与其长期目标保持一致。

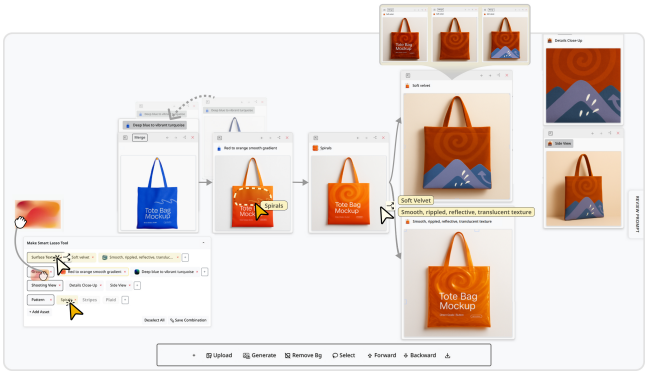


**图3：DesignTrace：一款用于探索备选方案并追踪设计进度的生成式人工智能设计工具。设计师可以从灵感图像和系统建议中提取语义属性，将其应用于选定区域或整幅图像，在共享语义类别下分支出多个备选方案，自定义图像生成提示词，在浏览编辑历史的同时进行局部编辑，并收起未使用的版本。**

## 3 重新思考人工智能的角色：将对齐视为不断变化的伙伴关系

生成式人工智能（GenAI）迅速融入设计实践，这要求我们不再将人工智能仅仅视为被动工具，而应将其看作能够共享能动性的伙伴。正如 ImageSense [5] 所展示的，协作式人机协同构思环境不止于被动搜索——它依托伙伴关系蓬勃发展，双方共享能动性，角色也灵活流动。因此，我们不将“对齐”视为遵循人类指令的固定状态，而是将其看作对主动权的动态协商；主动权会根据设计师当前的认知状态和创作过程所处阶段，在主导与跟随之间转换 [10]。

在人际协作中，领导权很少是一成不变的；随着合作伙伴逐渐形成对问题的共同理解，领导权也会随之变化。同样，人工智能代理也必须能够支持不同程度的主动性，才能成为有效的伙伴。

### 3.1 主动式人工智能：通过富有成效的摩擦进行设计

对齐可能受益于“富有成效的摩擦”。借鉴“主动权转移”模型 [10]，人工智能可以承担领导角色，提出“极端建议”或作为灵感来源，从而拓展创意空间 [5]。尽管高主动性的代理有时可能被认为不够“温暖”，但当用户缺乏明确路径时，他们凭借效率、决断力以及推动流程向前发展的能力而受到重视 [10]。主动式人工智能与人类互动中的对齐，关注的是人工智能理解设计空间的广度，并主动激发发散，而非输出的精确性。

### 3.2 响应式人工智能：以工具性意图进行设计

当设计师具有明确的工具性意图时，人工智能可能转而承担辅助性的“跟随者”角色。当设计师知道自己在寻找什么时，他们更偏好能够保留控制权的工具，例如语义搜索或人类协作，而不是主动式人工智能的建议。与人类计划保持行动对齐的“跟随者”代理被认为更加温暖、更有能力，也促进了与团队更强的社会认同。当人工智能试图在执行密集型阶段采取主动时，可能会被认为“固执”，或扰乱设计师的控制感。因此，响应式人工智能与人类互动中的对齐要求系统识别何时应将控制权交还给设计师，以确保信任和用户满意度。

### 3.3 设计实践中的主动权转移

在我们的研究中，设计师重视能够根据当下需求，在偶发性探索与借助人工智能系统专注工作之间灵活切换的能力 [5]。同样，人工智能伙伴必须能够判断互动情境：知道何时通过发挥主导作用带来“富有成效的摩擦”，何时严格遵循设计师的意图。这种主动权的灵活转移映照了人际协作的自然动态，并在整个创作过程中促进有意义的对齐。

## 4 结论

我们认为，设计师与生成式人工智能之间的对齐并非通过更可控的提示词和输出所实现的静态目标，而是在创作过程中动态展开的多阶段、多模式协同。创意实践并不要求我们仅在单一交互节点上优化对齐——主要通过改进提示机制或让输出更易于引导——而是要求实现异质对齐。通过提出三种交互方法，我们展示了如何通过不同的协商模式持续构建并重构对齐：可配置多模态脚手架通过组织设计师的输入来支持对齐；可直接操控的视觉提示通过直接的视觉协商促成对齐；可追踪的迭代工作流则通过将设计依据具体化并支持过程反思来促进对齐。

所提出的多阶段、多模式交互方法，契合创意工作的情境性与非线性特征——需求在实践中逐渐浮现，创意通过不断重组而逐步完善，而对齐本身也成为创意对话的一部分。这种多阶段视角重新定位了人工智能：它是一个协作伙伴，其主动性和响应能力可以在创意探索的不同阶段动态协调。我们呼吁未来的生成式人工智能系统能够在主动与响应角色之间灵活切换——有时激发发散思维以拓展设计空间，有时精准执行经过完善的意图，有时则退居幕后，以维护设计师的自主性。我们的目标是在整个工作流中分配结构与协调机制，从而支持创意实践中不断演进、互惠协商的内在过程。

# 参考文献

[1] Yining Cao、Yiyi Huang、Anh Truong、Hijung Valentina Shin 和 Haijun Xia。2025年。将组合结构作为人机共创环境的基础：一种设计方法与案例研究。载于 *2025*年*CHI*计算机系统人因会议论文集（*CHI '25*）。计算机协会，美国纽约州纽约市，文章 188，共 25 页。doi:10.1145/3706598. 3713401[2] DaEun Choi、Sumin Hong、Jeongeon Park、John Joon Young Chung 和 Juho Kim。2024年。CreativeConnect：借助生成式人工智能支持平面设计构思中的参考素材重组。载于 *2024*年*CHI*计算机系统人因会议论文集（美国夏威夷州檀香山）（*CHI '24*）。计算机协会，美国纽约州纽约市，文章 1055，共 25 页。doi:10. 1145/3613904.3642794[3] Hai Dang、Lukas Mecke 和 Daniel Buschek。2022年。GANSlider：用户如何通过多个滑块（有无前馈信息）控制图像生成模型。载于 *2022*年*CHI*计算机系统人因会议论文集（美国路易斯安那州新奥尔良）（*CHI '22*）。计算机协会，美国纽约州纽约市，文章 569，共 15 页。doi:10.1145/3491102.3502141[4] Nanna Inie 和 Peter Dalsgaard。2020年。交互设计师如何使用工具管理创意。*ACM* 人机交互汇刊27，2，文章 7（2020年3月），26 10 1145 3365104 页。doi: . / Janin Koch、Nicolas Taffin、[5] Michel Beaudouin-Lafon、Markku Laine、Andrés Lucero 和 Wendy E. Mackay。2020年。ImageSense：支持多样化人机合作关系的智能协作构思工具。*ACM*人机交互会议论文集4，计算机支持的协同工作1（2020年5月），45:1–45:27。doi:10.1145/3392850[6] Janin Koch、Nicolas Taffin、Andrés Lucero 和 Wendy E. Mackay。2020 年。SemanticCollage：利用语义标签丰富数字情绪板设计。载于 *2020*年*ACM*交互系统设计会议论文集（*DIS '20*）。计算机协会，美国纽约州纽约市，407–418。doi:10.1145/3357236.3395494[7] Rafal Kocielnik、Saleema Amershi 和 Paul N. Bennett。2019年。你会接受不完美的人工智能吗？探索调整终端用户对人工智能系统期望的设计。载于 *2019*年*CHI*计算机系统人因会议论文集（英国苏格兰格拉斯哥）（*CHI '19*）。计算机协会，美国纽约州纽约市，1 14 10 1145 3290605 3300641 – 。doi: . / . David [8] Chuan-En Lin、Hyeonsu B. Kang、Nikolas Martelaro、Aniket Kittur、Yan-Ying Chen 和 Matthew K. Hong。2025年。Inkspire：通过类比式草图绘制，借助生成式人工智能支持设计探索。载于 *2025*年*CHI*计算机系统人因会议论文集（*CHI '25*）。计算机协会，美国纽约州纽约市，文章 427，共 18 页。doi:10.1145/ 3706598.3713397[9] Zichen Liu、Yue Yu、Hao Ouyang、Qiuyu Wang、Ka Leong Cheng、Wen Wang、Zhiheng Liu、Qifeng Chen 和 Yujun Shen。2025年。MagicQuill：智能交互式图像编辑系统。载于 *2025*年*IEEE/CVF* 计算机视觉与模式识别会议（*CVPR*）。IEEE，美国田纳西州纳什维尔，13072–13082。doi:10.1109/CVPR52734.2025.01220[10] Inês Lobo、Janin Koch、Jennifer Renoux、Inês Batina 和 Rui Prada。2024年。何时应当主导或跟随：理解人机协作游戏中的主动性水平。载于 *2024*年*ACM*交互系统设计会议论文集（丹麦哥本哈根）（*DIS '24*）。计算机协会，美国纽约州纽约市，2037 2056 10 1145 3643834 3661583 – 。doi: . / . [11] 1993 Brian Logan 和 Tim Smithers。创造力与设计即探索。载于 创造力建模与基于知识的创意设计。心理学出版社，希尔斯代尔，新泽西州。页数：37。[12] Yuwen Lu、Alan Leung、Amanda Swearngin、Jeffrey Nichols 和 Titus Barik。2025年。Misty：通过交互式概念融合进行用户界面原型设计。载于 *2025*年*CHI*计算机系统人因会议论文集（*CHI '25*）。计算机协会，美国纽约州纽约市，文章 1108，共 17 页。doi:10.1145/3706598.3713924[13] Hyerim Park、Joscha Eirich、Andre Luckow 和 Michael Sedlmair。2024年。“我们是视觉思维者，而非语言思维者！”：专业设计师如何使用生成式人工智能图像生成工具的主题分析。载于 第*13*届北欧人机交互会议论文集（瑞典乌普萨拉）（*NordiCHI '24*35，14 10 1145 3679318）。计算机协会，美国纽约州纽约市，文章页码。doi: . / . 3685370 [14] Xiaohan Peng、Debanjana Haldar、Janin Koch 和 Wendy E. Mackay。2026年。DesignTrace：借助生成式人工智能探索、迭代和追踪设计方案。载于*2026*年*CHI*计算机系统人因会议论文集（西班牙巴塞罗那）（*CHI '26*）。计算机协会，美国纽约州纽约市。doi: 10 1145 3772318 3791036 . / . Xiaohan Peng、Janin Koch 和 [15] Wendy E. Mackay。2024年。DesignPrompt：运用多模态交互，通过生成式人工智能进行设计探索。载于 *2024*年*ACM*交互系统设计会议论文集（丹麦哥本哈根）（*DIS '24*）。计算机协会，美国纽约州纽约市，804–818。doi:10.1145/3643834.3661588

[16] Xiaohan Peng、Janin Koch 和 Wendy E. Mackay。2025年。FusAIn：运用笔式交互编写生成式人工智能视觉提示词。载于*2025*年*CHI*计算机系统人因会议论文集（*CHI '25*）。计算机协会，美国纽约州纽约市，文章874，20页。doi:10.1145/3706598.3714027[17] Vishnu Sarukkai、Lu Yuan、Mia Tang、Maneesh Agrawala 和 Kayvon Fatahalian。2024年。区块与细节：为草图生成图像提供支架。载于第*37*届*ACM*用户界面软件与技术研讨会年会论文集（美国宾夕法尼亚州匹兹堡）（*UIST '24*）。计算机协会，美国纽约州纽约市，文章 33, 13 10 1145 3654777 3676444 页。doi: . / . [18] 2017 Donald A. Schön。《style id='10'>反思性实践者》。劳特利奇，英国阿宾顿。doi:10.4324/9781315237473[19] Donald A. Schon 和 Glenn Wiggins。1992年。设计中的观看方式及其作用。《设计研究》13卷，2期（1992年），135–156页。doi:10.1016/0142-694X(92)90268-F [20] Kihoon Son、DaEun Choi、Tae Soo Kim 和 Juho Kim。2024年。揭开平面设计中隐性知识的神秘面纱：特征、实例、方法与指南。载于*2024*年*CHI*计算机系统人因会议论文集（*CHI '24*）。计算机协会，美国纽约州纽约市，1 18 10 1145 3613904 3642886 –。doi: . / . Sirui Tao、Ivan [21] Liang、Cindy Peng、Zhiqing Wang、Srishti Palani 和 Steven P. Dow。2025年。DesignWeaver：为文生图产品设计提供维度支架。载于*2025*年*CHI*计算机系统人因会议论文集（*CHI '25*）。计算机协会，美国纽约州纽约市，文章 425, 26 10 1145 3706598 3714211 页。doi: . / . Veera Vimpari、[22] Annakaisa Kultima、Perttu Hämäläinen 和 Christian Guckels-berger。2023年。“一种不进则退的处境”：游戏行业专业人士对文生图人工智能的感知、采纳与使用。*ACM*人机交互论文集7卷，CHI PLAY，文章379（2023年10月），34页。doi:10.1145/3611025[23] Qinyi Zhou、Jie Deng、Yu Liu、Yun Wang、Yan Xia、Yang Ou、Zhicong Lu、Sai Ma、Scarlett Li 和 Yingqing Xu。2025年。ProductMeta：一种利用多模态大语言模型进行隐喻式产品设计构思的交互系统。载于*2025*年*CHI*计算机系统人因会议论文集（*CHI '25*）。计算机协会，美国纽约州纽约市，文章428，24页。doi:10.1145/3706598.3713935