arXiv:2601.00475v1 [cs.AI] 1 Jan 2026

利用智能体人工智能框架进行渐进式构思，实现人机共同创造

B. Sankar*印度科学学院（IISc）机械工程系，印度班加罗尔
sankarb@iisc.ac.in

Srinidhi Ranjini Girish 印度科学学院（IISc）设计与制造系，印度班加罗尔

Aadya Bharti 印度科学学院（IISc）设计与制造系，印度班加罗尔

Dibakar Sen 印度科学学院（IISc）设计与制造系，印度班加罗尔

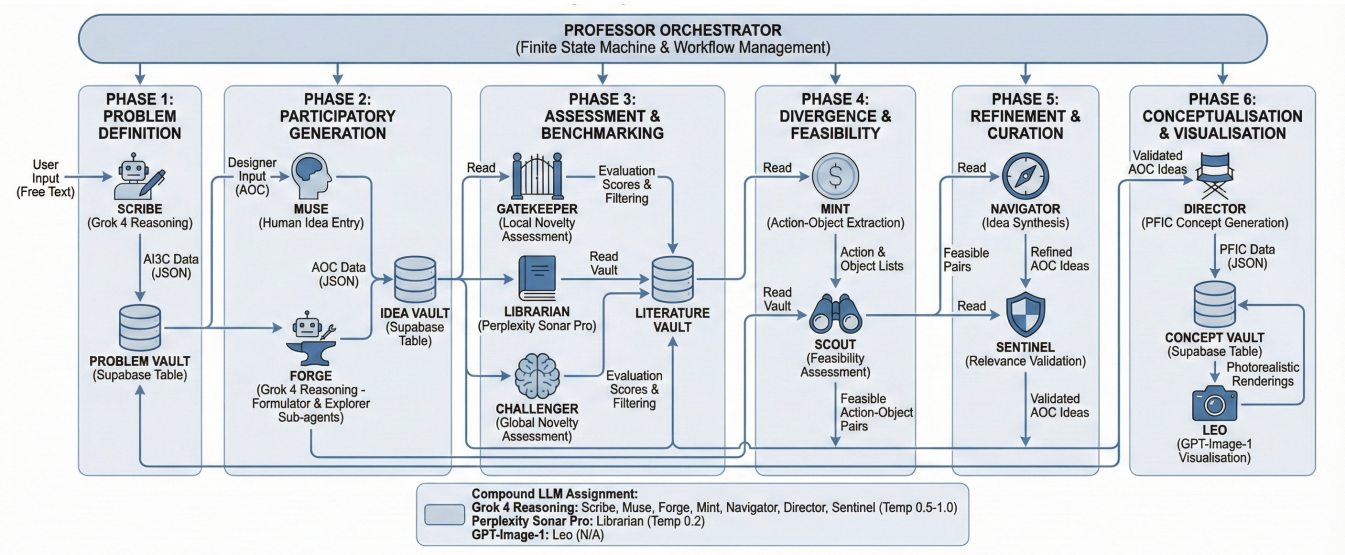


图 1：MIDAS（利用分布式智能体人工智能系统进行元认知构思）框架的人机共同创造环境。该插图展示了协作式界面：用户与多样化的专业化人工智能智能体群体进行交互，以促进渐进式构思。MIDAS 的分布式智能体架构展示了分布在六个运行阶段中的十三个专业化智能体，这些智能体通过持久化 Vaults 相互连接，并由 Professor 智能体进行编排。

摘要

对于当代工程设计而言，生成真正新颖且多样化的创意至关重要，但这对新手设计师而言仍是一项重大的认知挑战。当前“单次突发式”人工智能系统生成大量语义聚类的创意，进一步加剧了这一挑战。我们提出了 MIDAS（通过分布式智能体人工智能系统实现元认知构思），这是一种新颖的框架，以专门化人工智能智能体组成的分布式“团队”取代单一人工智能范式，旨在模拟人类的元认知构思工作流程。该智能体系统逐步完善创意，并分别根据全局新颖性（相对于现有解决方案）和局部新颖性（相对于先前生成的创意）对每个创意进行评估。因此，MIDAS 展示了一种可行且渐进式的真正人机共创范式，将人类设计师从被动筛选者提升为参与式、主动式的协作伙伴。

关键词

构思、产品设计、智能体人工智能、协作、人工智能

1 引言

构思是生成、发展和交流新颖创意的认知过程，也是工程设计过程中至关重要的环节 [15]。这一阶段的成效取决于人类设计师的内在专业能力、领域知识和默会经验。传统构思方法，如头脑风暴 [20]，综摄法 [21]，TRIZ [9, 22]，等，旨在提供程序性支持；然而，根本的创造驱动力仍源自人类认知。这

*通讯作者

依赖个人或群体智力本身也有局限，例如认知偏差、设计固着 [24]，群体环境中的产出阻滞，以及任何单个人都不可能真正拥有涵盖众多不同领域的百科全书式知识。

在现代产品设计复杂且多学科交织的情境中，尤其需要快速生成大量新颖且多样化的创意，因此有必要探索计算支持 [19, 23]。借助机器智能增强或辅助人类创造力并非新鲜事；数十年来，人们一直以各种术语探讨这一课题（图2），例如计算机辅助构思（*CAI*）、人机协同构思、协同构思、人工智能增强构思、共同创造、人机协作、参与式设计、计算创造力、混合主导式构思等 [7, 16]。如图 2 所示，这些术语上的多样性反映出学术讨论的核心在于人与计算机之间的关系 [7, 16]。

无论采用何种具体术语，这种协作关系中机器的基本作用都是减轻人类设计师的创意负担。早期的计算机辅助构思工具只是被动的信息库或简单的辅助工具，用于处理形态分析或组合探索。创意负担仍由人类设计师承担。自 2020 年代以来，生成式人工智能，尤其是大型语言模型（LLM）的指数级发展与普及，展现出根据输入的问题陈述 [1, 8, 17, 26–28, 31]在瞬间生成大量创意的潜力。机器已从被动的工具转变为主动的协作者，能够生成内容，其成果往往与人类创作难以区分，有时甚至优于人类 [6, 12, 17, 25, 31]。

然而，创意数量过于庞大，使人类专家在评估其质量时面临挑战，因为这一过程耗费资源、对认知能力要求高，并且难以保持评估标准的一致性。早期研究已证明，可以利用向量嵌入和聚类来完成这项评估 [18]。研究表明，人工智能生成的创意往往围绕少数几个基础概念形成语义簇，无法提供设计师所需要的真正新颖性和多样性 [18]。本文的主要目标是通过提出一个能够生成大量真正新颖创意的系统来弥合这一差距。我们对新颖性作出双重定义：创意必须不同于文献中已有的解决方案（全局新颖性），同时也必须不同于同一创意会话中生成的其他创意（局部新颖性）。我们认为，无论是人类构思还是计算机构思，所有构思都会按照时间顺序展开。因此，$(n+1)^{th}$ 创意不仅必须相对于文献（现有解决方案）具有新颖性，还必须相对于在其之前生成的 n 创意具有新颖性。这一时间感知要求是现有基于人工智能的构思系统的主要局限，它们采用“单次突发式”模型，瞬间生成一大批创意。我们提出利用这一局限。在本文中，我们提出、设计并实现了一种基于智能体人工智能系统的新框架，该框架经过明确架构设计，旨在管理构思中的这一渐进式、时间连续且持续的生成—评估工作流。

2 渐进式构思的潜力

真正由大语言模型驱动的构思正受到单次突发式、单体化系统所引发的诸多问题的阻碍。要构建一个能够真正与设计师协作的系统，我们必须首先明确并解决这些关键缺陷。

2.1 当前人工智能辅助构思中的“单次突发式”局限

当前探索大语言模型在设计中应用的研究正蓬勃发展 [3]。然而，主流交互模式仍然简单。大多数研究和商业工具都将人机协作设定为一种二元对话：设计师与单一、整体式的大语言模型（如 GPT、Grok 或 Gemini 等）进行轮流对话 [4, 10]。这种交互常被类比为与一位伙伴或朋友进行头脑风暴，而对方拥有极为广博的知识库，能够迅速关联彼此迥异的概念，从而激发人类设计师自身的创意思路。

尽管这种对话式范式已被证明有助于基础的创意生成，并且研究显示它能够增加创意的数量 [17]，，但它存在一个根本缺陷：它依赖于单次突发式的查询-响应模型。这种交互是交易式的；设计师提供提示词，大语言模型随即返回一阵“突发式”的创意。这种模型过度简化，无法体现真实人类构思错综复杂、反思性且渐进式 的本质。

人类设计师不会仅凭一闪而过的灵感就生成最终且稳健的概念。这一过程本质上是迭代式、杂乱且渐进的。它涉及复杂的元认知活动交互：定义问题、生成初步想法、根据约束对其进行评估、研究外部刺激、基于新洞见重新构建问题框架，并通过多个步骤逐步将一个“模糊”的想法完善为具体且可采取行动的解决方案。当前将单个大语言模型作为合作伙伴的模型无法复现这一结构化、多阶段的认知历程。它更像一台通用的“答案机器”，而非专门的设计系统，因此存在多项严重缺陷，阻碍了真正新颖且具有影响力的解决方案的生成 [5, 13]。

2.2 从群体动态到单体模型

在专业设计实践中，构思很少是单打独斗的。人们通常会以小组形式开展构思，以发挥成员的认知多样性。成功的人类头脑风暴团队由多人组成，成员彼此协作、相互批评，并积极参与，共同提出一系列可行的解决方案 [20]。团队中的每位成员都接受过不受限制的创意生成、批判性分析和可行性评估方面的训练，能够调研现有技术，并将综合不同的创意，形成连贯的整体。集体活动中的这种动态互动，正是推动创新的动力 [11, 14]。当前的大语言模型舍弃了这种丰富的动态过程，取而代之的是一个通才型的单一“伙伴”。尽管人类群体可能会遭遇权威偏见或产出阻滞等社会认知问题，但解决之道并非消除群体互动，而是对其加以优化。我们认为，还存在一种强大得多的范式：一种

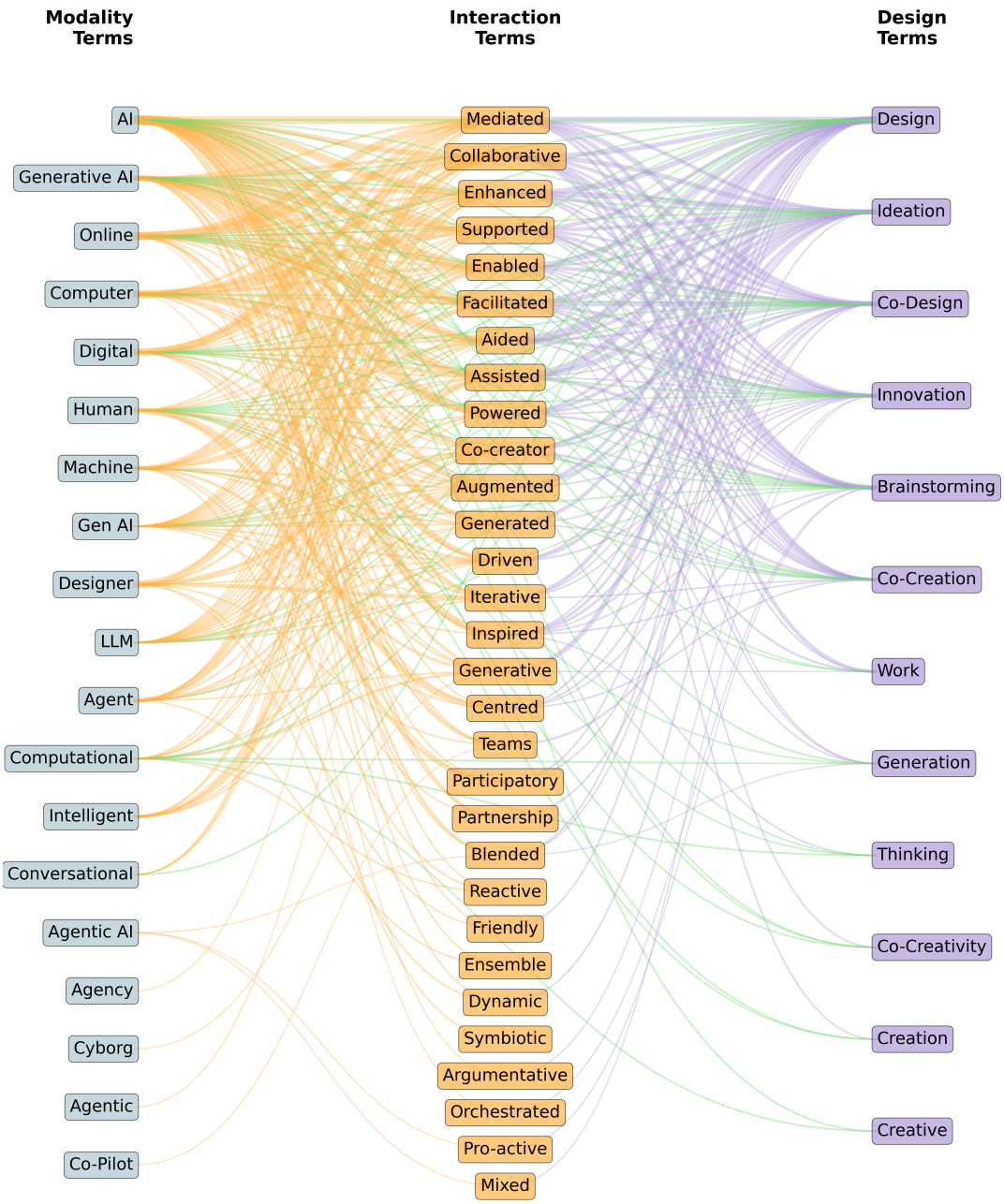


图 2：桑基图展示了人工智能辅助创作流程中模态术语（如人工智能、生成式人工智能、大语言模型）、交互术语（如协作式、增强式、共生式）与设计术语（如构思、协同设计、创新）之间的关联。

由一位人类设计师和一支由多个专业化人工智能智能体组成的团队构成的混合群体。在这样的系统中，每个智能体都针对特定功能进行配置。这正是智能体人工智能系统的基础；它从简单的二元伙伴关系转向分布式多智能体协作结构，模拟高绩效的人类设计团队。

2.3 阶段关卡与持续生成及评估

大多数人工智能构思工具优先考虑生成，产出大量创意 [17], ，却缺乏评估支持。设计师只能手动筛选创意的新颖性和可行性，这只是转移了负担，而非共同创作。虽然已有一些机器评估相关研究，但

这些研究通常在生成完成后独立使用。在软件工程领域，持续集成/持续部署（CI/CD）的出现，通过在编写、测试和部署代码之间建立紧密的反馈循环，彻底改变了开发方式。我们认为，设计构思也需要类似的模式：一条*CG/CA*（持续生成／持续评估）流水线。在这种模式下，生成和评估并非彼此独立、依次进行的阶段，而是在持续循环中深度交织。创意生成后会立即接受新颖性和多样性评估，再根据评估结果为下一轮生成提供动力，如此循环往复。这种迭代优化是确保创意池逐步完善的唯一方法。此前的研究已经为此类评估建立了稳健的计算框架，利用向量嵌入和聚类，通过创意稀疏度和簇稀疏度等指标量化新颖性和多样性 [18], ，为流水线中的“CA”部分奠定了方法基础。

2.4 “新颖性幻觉”：幻觉、变体与固着

当前大语言模型一个重要且常未被报告的问题，是其生成的“新颖性”的质量。许多研究声称，大语言模型生成的创意比人类更具新颖性，但这一说法往往基于有偏差的样本。大语言模型从本质上说是随机鹦鹉，基于现有人类知识的中位数进行训练。它们擅长插值，也就是创造听起来合理、但只是现有解决方案变体的内容；然而，它们难以实现真正的外推，即激进且打破常规的创新。这导致了“千篇一律的海洋”：数以千计生成的“创意”不过是彼此在句法或语义上的细微变化；我们将这种现象称为“新颖性幻觉”。

缺乏现实世界依据进一步加剧了这一问题。现有基于人工智能的构思系统的另一缺陷是，它们会生成创意，却没有机制将其与大量现有解决方案进行比对，例如专利、学术论文和商业产品。某个创意对设计师和大语言模型而言可能显得新颖，但它或许早在十年前就已获得专利。因此，我们的目标并非追求“创意丰沛度”，而是为设计师提供少量真正新颖且多样化的创意——无论是彼此之间（局部新颖性），还是相对于现有解决方案而言（全局新颖性）。

2.5 线性构思：尚未发掘的「创意源自创意」潜力

当前人工智能构思的主流模式是线性且单向的：设计师输入一个问题（P1），人工智能便生成一系列创意 {I1, I2, I3...I5}。这种模式假设创意与其产生时所针对的问题紧密相连。我们认为，这种方法未能捕捉人类创造力的核心机制：建立不同事物之间的联系、概念融合和类比迁移。创意并非天生局限于某个单一问题。针对问题 P1 生成的一个创意（例如 I4），也可能是另一个完全不同的问题（P2）的可行解决方案。而这个新问题 P2 也会有自己的一系列现有解决方案 {I4, I5, I6, I7, I8}。真正深刻的创造性飞跃在于进一步追问：这些针对 P2 的其他解决方案（例如 I6 或 I7）能否调整后用于解决原问题 P1？我们认为，如图 ref fig:nonlinear-ideation-network 所示，这种生成“由创意衍生创意”的非线性网络化方法，能让设计师突破最初问题的概念边界，而这是当前线性生成模型完全不具备的能力。

2.6 「评估过载」与设计师被动角色

设计师往往不倾向于在创作过程中使用人工智能；这并非卢德主义，而是对一种主动贬低其贡献的范式所作出的反应。这是因为这些系统没有提供整合设计师自身创意的机制。人类的创意火花被搁置一旁，导致参与度下降、归属感丧失。这是采用此类系统的一种心理障碍。为克服这一问题，我们提出了「*PAC*」（参与式、主动式、协作式）构思框架，详见第 3.2 节。在基于 PAC 的系统中，设计师居于核心，而非处于边缘。他们是主动参与者，也会向系统贡献自己的创意，而不只是提供提示。随后，这些人类生成的创意会与机器生成的创意一同进入评估与开发流水线。评估采用盲审方式，评审者并不知道或不在意创意源自人类还是人工智能；所有创意都按其自身价值接受评估。这将鼓励设计师保持参与，并建立真正的协同伙伴关系。

2.7 渐进式构思的理论依据：实验发现

为验证本文的核心假设，我们首先开展了一项简短实验，该实验基于此前发表的研究。此前的论文 [17] 介绍了 DesignerAI，这是一种经过微调的大语言模型，能够通过单次突发式生成“大量”创意。研究在六个截然不同的问题陈述（PS1 至 PS6）中测试了这一能力，涵盖广泛的领域。

- **PS1：** 用于分类的产品，以实现有效的废弃物管理。
- **PS2：** 用于鞋类消毒和清洁的产品，以改善卫生状况并提升安全性。
- **PS3：** 用于提高家用餐具清洁效率和可持续性的产品。
- **PS4：** 用于提高排队队列中长时间站立舒适度和效率的产品。
- **PS5:** 用于在家喂鸟、促进老年人心理健康的产品。
- **PS6:** 便于旅行时晾干和收纳雨伞的产品。

对于每个问题陈述，DesignerAI 都生成了 100 个创意。在后续研究中， [18], 研究人员使用一种基于向量嵌入和聚类的新框架，对这些创意集进行了计算评估。图4、图5和图6左列展示了这些领域中具有代表性的聚类图。如可视化图像所示，在全部六个领域中，100 个创意都汇聚成了不同的高密度簇。每个簇中的创意

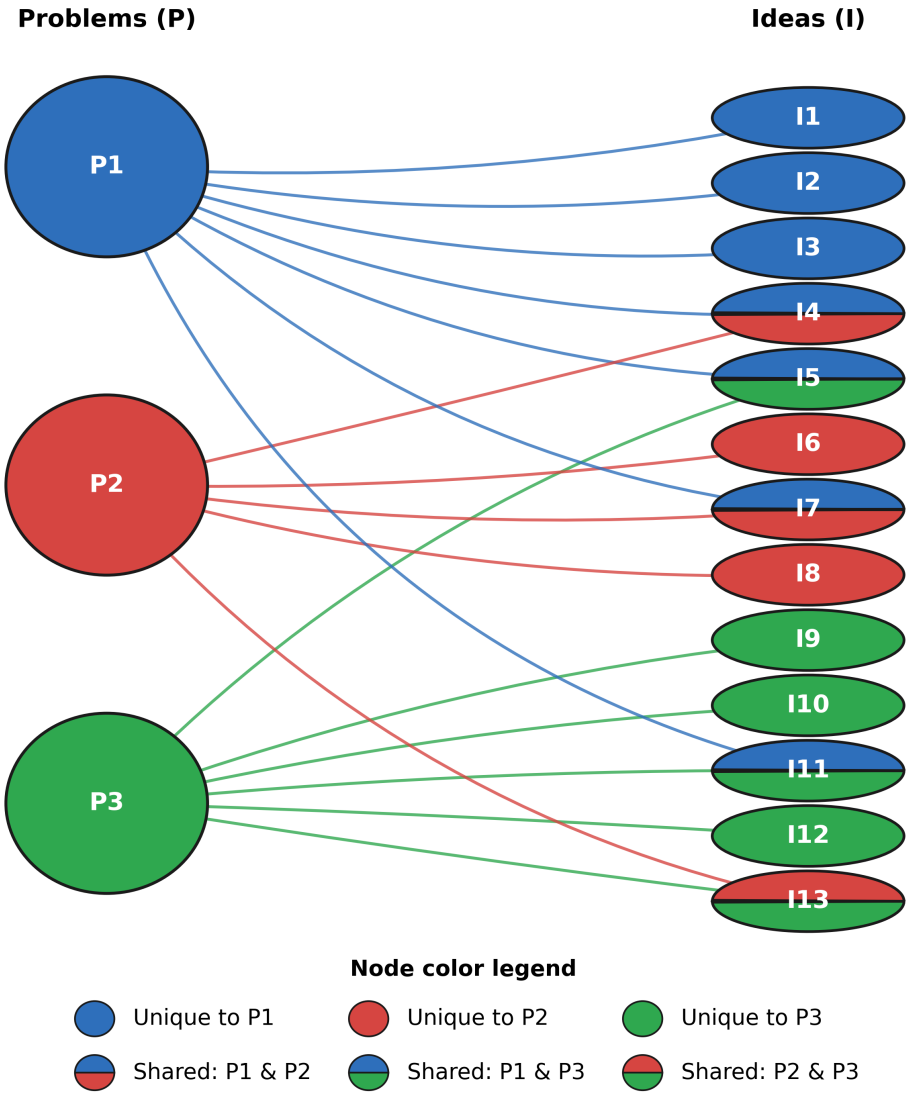
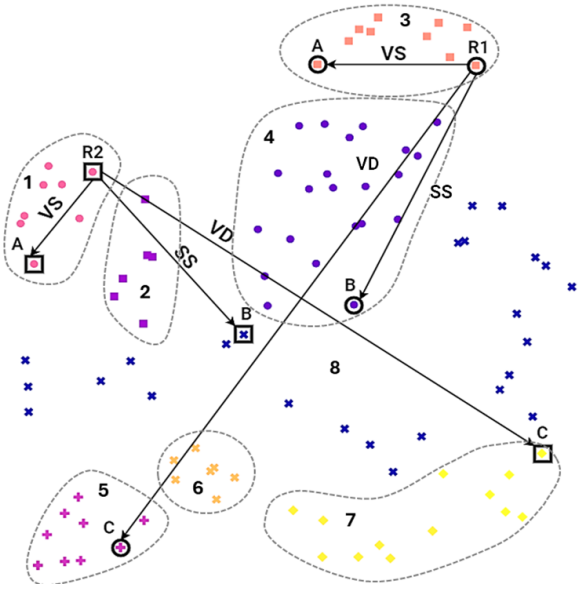


图 3：通过“由想法衍生想法”实现的网络化非线性构思。与主流的线性模型不同，在该模型中，一个问题 P_1 会产生一份单向排列的创意列表 $\{I_1, \ldots\ldots\}$；此图将构思描绘为一个可复用、跨问题的网络。问题（P）显示在左侧，创意节点（I）显示在右侧，边表示适用关系。关键在于，针对某个问题生成的一些创意也适用于其他问题（例如，I_4 关联 P_1 和 P_2；I_5 关联 P_1 和 P_3），从而促成概念融合与类比迁移。这种“重新关联”构建出一条条路径，使得在 P_2 下发现的创意（例如，I_6–I_8）能够调整后用于解决 P_1（例如，I_7），其他问题与创意的组合也同样如此（例如，I_{13} 关联 P_2 和 P_3，而 I_{11} 关联 P_1 和 P_3）。节点颜色表示创意是某个问题独有还是多个问题共享，凸显促成非线性创意跃迁的桥梁。

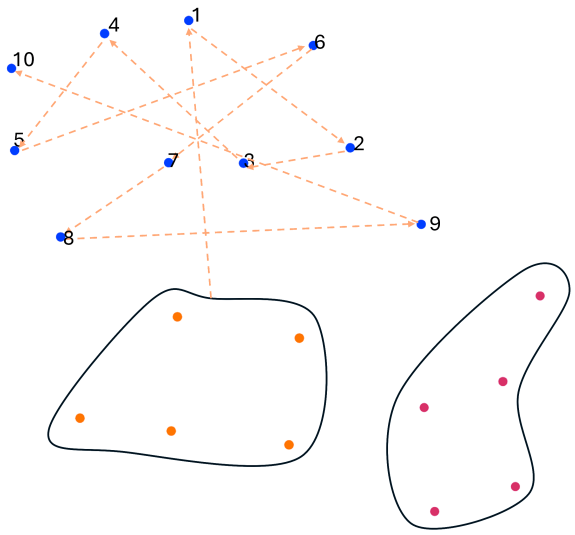
在语义上都非常相似（VS），而簇与簇之间则截然不同（VD）。这从经验层面揭示了单次突发式生成的一个关键局限：基于大语言模型的构思虽然能产出大量创意，但其中一些创意实际上只是语义变体，紧密围绕少数几个核心新颖创意展开。

随后，我们试图确定能否“激发”大语言模型克服这一局限，逐步生成创意。对于每个

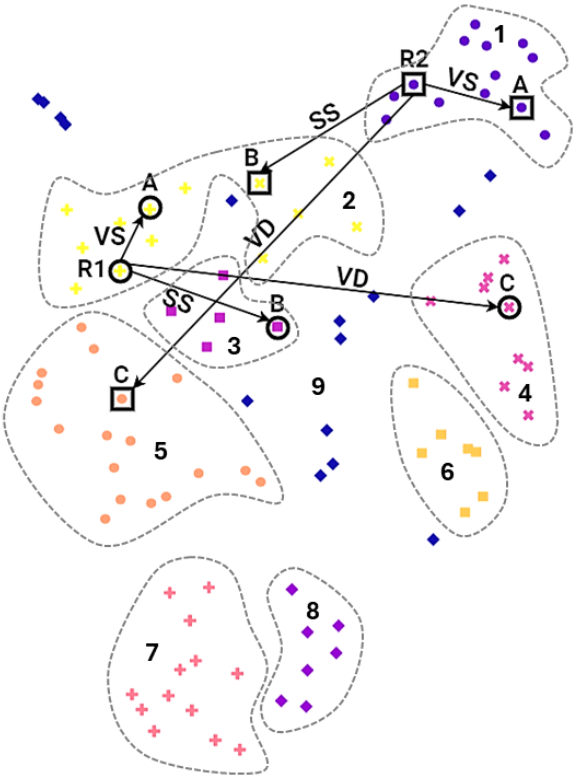
六个问题陈述，我们首先提示 DesignerAI 一次性生成 10 个创意。随后，我们通过迭代提示模型，在接下来的 10 轮中每轮生成一个新创意。每一步都施加了一项关键约束：新创意必须在语义上区别于此前生成的所有创意（最初的 10 个以及后续逐步生成的所有创意）；这些创意都存储在其上下文缓冲记忆中，以供调用。



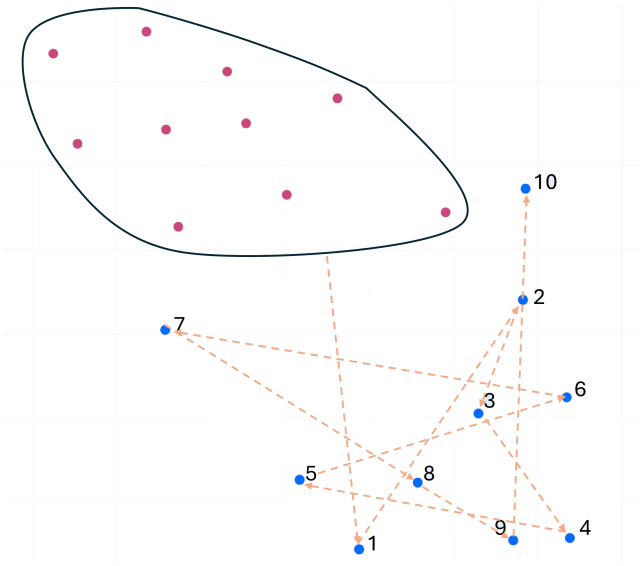
(a) PS1：单次爆发式聚类



(b) PS1：渐进式激发

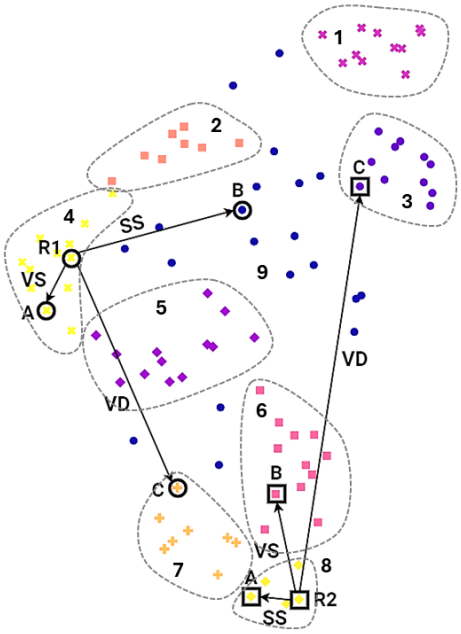


(c) PS2：单次爆发式聚类

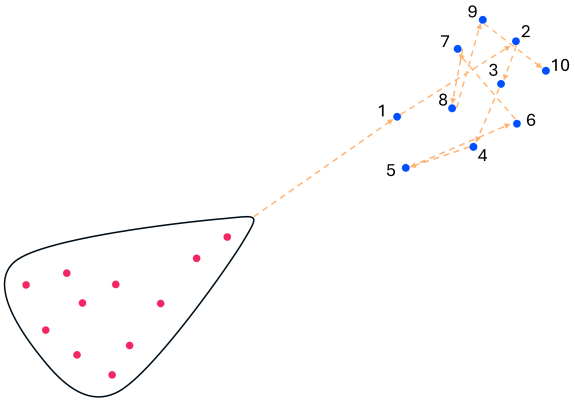


(d) PS2：渐进式激发

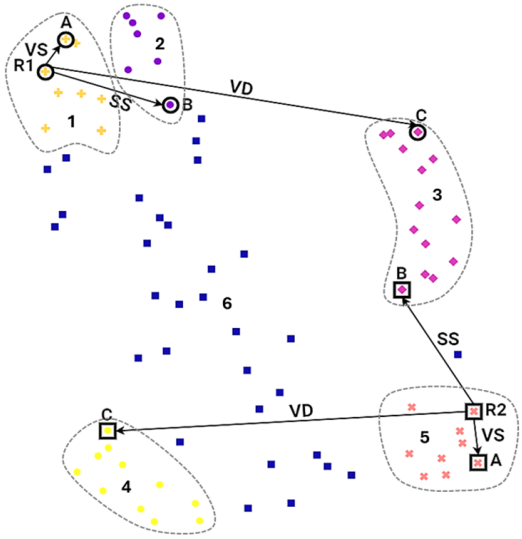
图 4：问题陈述 PS1（垃圾分类）和 PS2（鞋类消毒）中创意生成方法的语义聚类对比可视化。(a)、(c)“单次突发式”生成会形成相似创意的高密度簇。(b)、(d)“渐进式激发”促使生成语义距离较远且多样化的创意（蓝点）。在两个领域中，从基线的高密度簇转变为渐进式集合中的多样化微簇，这一变化都显而易见。



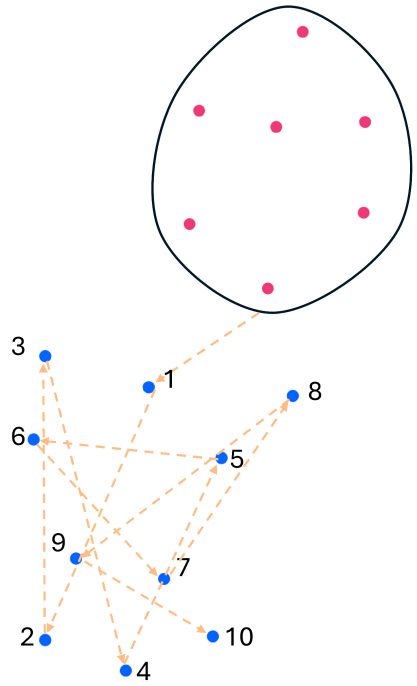
(a) PS3：单次爆发式聚类



(b) PS3：渐进式激发

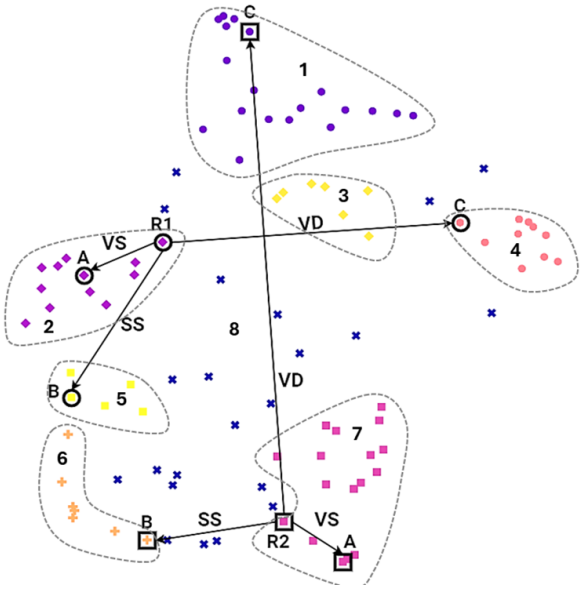


(c) PS4：单次爆发式聚类

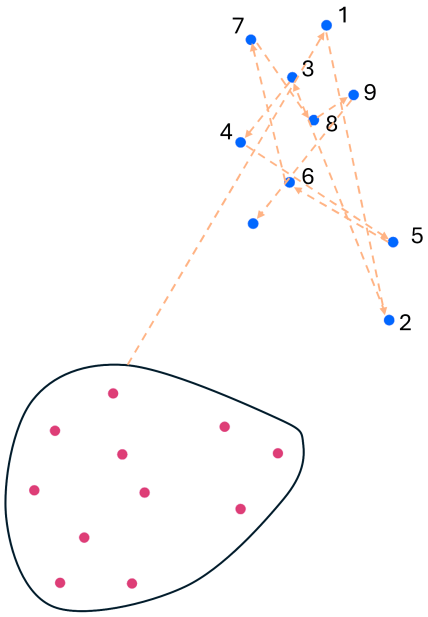


(d) PS4：渐进式激发

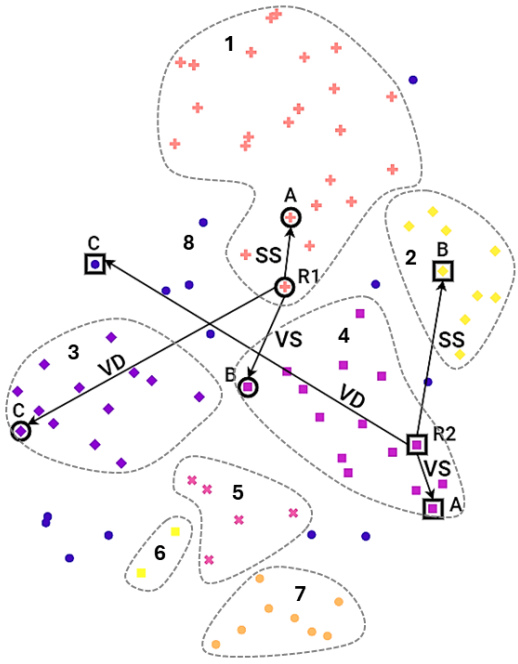
图5：问题陈述 PS3（洗碗）和 PS4（排队舒适度）下各类创意生成方法的语义聚类对比可视化。(a)、(c)“单次突发式”生成会形成包含相似创意的高密度簇。(b)、(d)“渐进式激发”促使生成语义距离较远且多样化的创意（蓝点）。在这两个领域中，从基线的高密度簇到渐进式集合中的多样化微簇，其转变都十分明显。



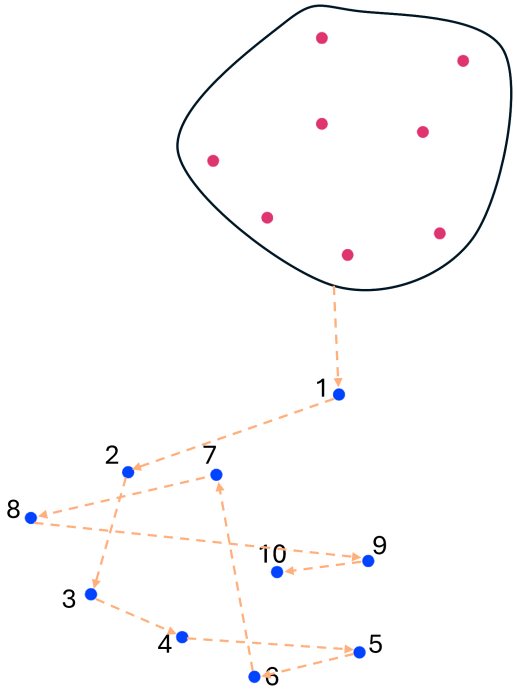
(a) PS5：单次爆发式聚类



(b) PS5：渐进式激发



(c) PS6：单次爆发式聚类



(d) PS6：渐进式激发

图6：针对问题陈述 PS5（喂鸟）和 PS6（雨伞收纳）的创意生成方法语义聚类对比可视化。(a)、(c)“单次突发式”生成会形成由相似创意构成的高密度簇。(b)、(d)“渐进式激发”促使生成语义距离较远且多样化的创意（蓝点）。在这两个领域中，从基线中的高密度簇过渡到渐进式集合中的多样化微簇，均清晰可见。

随后，针对每个 PS 的总计 20 个创意（10 个单次突发式 + 10 渐进式创意），使用
[18], 4, 中的相同
5 6
聚类算法进行处理，结果显示在图和图的右列中。在所有情况下，图表都清晰地显示，初始创意（以成组的粉色或橙色圆点表示）形成了密集的聚类，从而验证了我们关于单次突发式生成的假设。与此形成鲜明对比的是，10 个渐进式生成的创意（蓝色圆点，编号 1 至 10）始终与初始聚类保持显著的语义距离，每个创意都形成了自身独立的微簇。

尽管这些结果表明，通过迭代式激发，大语言模型具备生成真正新颖且多样化创意的潜在能力，但这种基于提示的引导也并非没有局限。仅依赖手动提示不仅劳动密集，而且缺乏结构化依据；它依赖模型内部的“回忆”，却没有正式机制根据全球现有技术或技术可行性进行基准评估。此外，随着迭代次数增加，简单提示往往难以维持语义距离，进而导致“幻觉式新颖性”或重复变体。这项实验为我们的研究提供了基础动机，表明渐进式构思虽然可行，但需要一个正式的多智能体框架，以便在可扩展且专业的设计情境中，系统地管理创意的生成、评估与完善。因此，本文旨在通过引入渐进式构思 这一概念，借助智能体人工智能框架弥合这一空白。

渐进式构思是一个渐进且迭代的过程，从模糊、定义不清的问题空间，逐步迈向一组具体、定义明确的解决方案。渐进式构思将“CG/CA”流水线付诸实践，确保每一轮生成的内容都比上一轮更新颖，并通过提供经过筛选的多样化创意，而非成百上千个变体，解决“过度丰 abundance问题”。因此，我们认为，当前的人机构思范式建立在单次突发式、二元和单体式大语言模型交互之上，存在缺陷。它未能体现人类创造力多阶段、重评估且渐进的特性，使人类设计师沦为被动的边缘角色。

我们的主要目标是：

(1) 设计并开发一种新颖的智能体人工智能架构，将构思这一复杂活动分解为由专业化协作智能体组成的分布式系统，突破单一大语言模型范式。(2) 在该架构中实现“PAC”（参与式、主动式、协作式）框架，将人类设计师置于核心位置，作为积极参与的共同创造者，使其贡献得到整合和重视。(3) 将“CG/CA”（持续生成／持续评估）流水线付诸实践，促进真正的渐进发展，逐步从模糊问题迈向具体概念。(4) 在该流水线中整合全局与局部新颖性检查，将生成的创意与现实世界的专利和产品进行比较，并相互比较，以筛选真正的创新，而非简单变体。

3 人机构思渐进式框架：MIDAS

为实现这些目标，我们提出了一个名为通过分布式智能体人工智能系统实现元认知构思（*MIDAS*）[1]的新型系统。MIDAS 是一个分布式多智能体系统，由 13 个不同的专业化智能体组成。这些智能体与人类设计师协同工作，管理从构建定义不明确的问题，到呈现最终筛选创意的照片级真实感概念的完整构思工作流程。为管理这些智能体之间复杂的信息流，该框架建立在一项关于设计交流中结构化知识表示的基础性前期研究之上 [17]。该研究确立了一种新型的设计产物分类法：用于表示问题陈述的*AI3C*（活动、项目、矛盾、标准、约束）模型、用于表示创意的*AOC*（行动、对象、情境）模型，以及用于表示概念的*PFIC*（原理、特征、实施、特性）模型。这种共享的结构化“语言”使专业化智能体能够无缝协作，以高保真度传递信息，并实现复杂的多阶段推理。

3.1 通过智能体模拟人类认知过程

人类构思，尤其是在复杂的工程设计中，并非单一事件，而是一个精密的认知过程，包含一系列不同的元认知活动，我们通过以下不同的智能体来体现这些活动：

- **Scribe** – 定义： 理解用户需求并综合洞见，以构建可执行、结构化的问题陈述。
- **Muse &Forge**– 头脑风暴与构思： 生成广泛的初步创意，并根据给定的问题陈述系统地生成解决方案。
- **Gatekeeper & Scout** – 评估： 根据约束条件评估生成的创意：新颖性和可行性。
- **Librarian 和 Challenger** –收集与比较： 将生成的解决方案与现有知识或现有技术进行比较。
- **Mint** – 发散： 拓展解空间的边界，从现有创意中发现语义新颖路径。
- **Navigator 和 Sentinel**– 优化与筛选： 通过类比推理和结构化优化来调整、融合并完善创意。
- **Director 和 Leo** –创作与构想： 将最终创意综合成连贯、详尽且可执行的概念提案和概念表现图。

MIDAS 的架构如图7所示，由上文列出的多个智能体组成一个分布式系统，这些智能体通过一系列持久化数据库相连，这些数据库称为“Vaults”。工作流程由一个编排器（称为“**Professor**”智能体）管理，负责指挥智能体之间的数据流动。该架构明确实现了**ReACT（推理 + 行动）**范式 [2, 29, 30]，，其中每个

[1]该系统可通过 https://midas-ai-ideation.vercel.app/ 在线访问。如需试用 MIDAS，请向我们发送请求以获取访问凭证

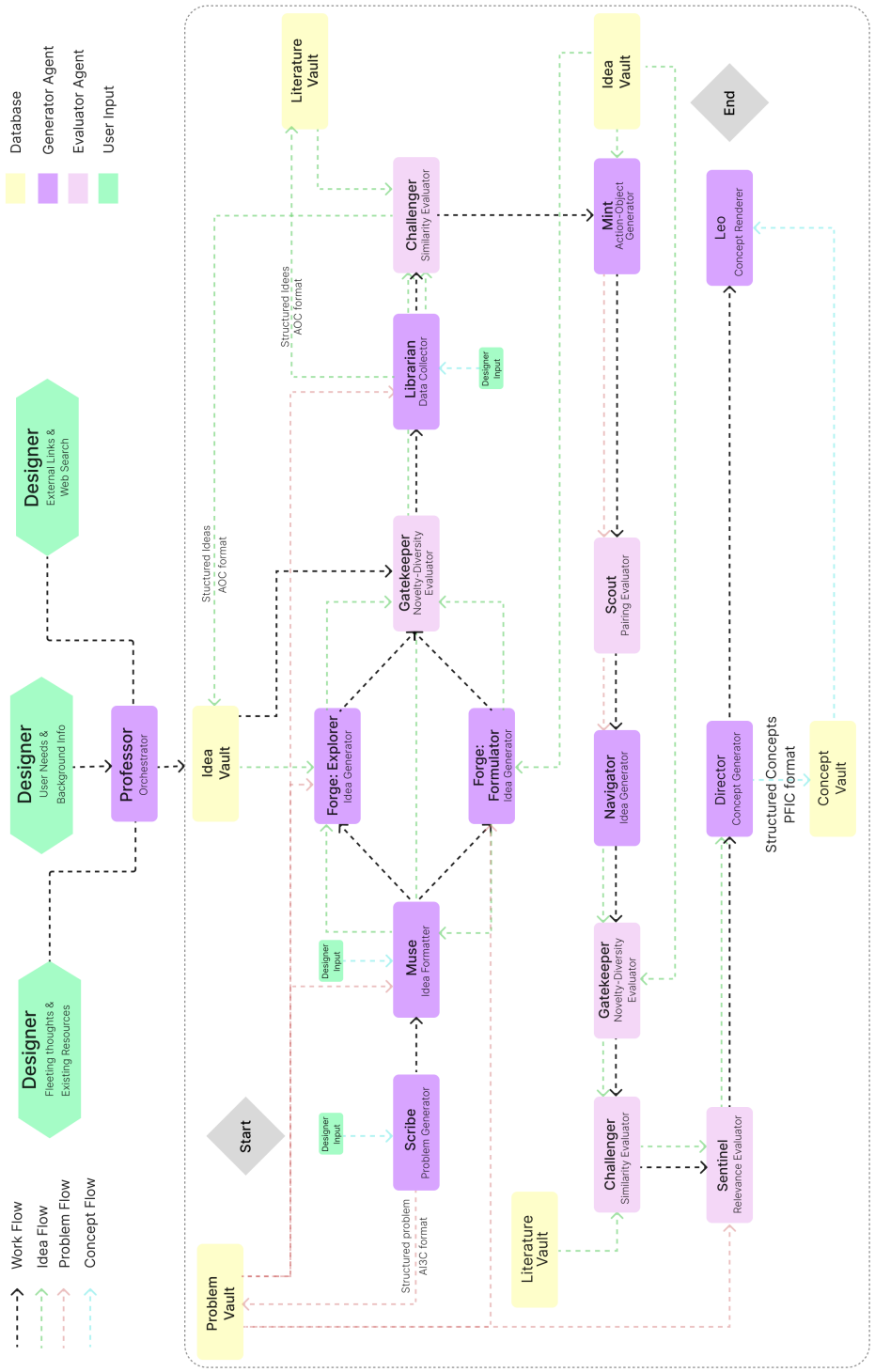


图7：MIDAS框架的详细系统架构，展示了工作、创意和概念在生成器智能体、评估器智能体、持久化 Vaults 与人类设计师之间的流动。

智能体的行动是更宏大认知过程中的一个离散且经过推理的步骤。

3.2 A'PAC'（参与式、主动式、协作式）构思伙伴关系

MIDAS架构的核心哲学原理之一，是建立一种**'PAC'（参与式、主动式、协作式）**伙伴关系，旨在克服设计师在传统人工智能工具中所排斥的"被动策展人"角色。MIDAS主要通过"Muse"智能体实现这一目标。

构思流程并非从人工智能生成解决方案开始。设计师将用户需求和问题提供给"Scribe"，由其生成结构化问题陈述；之后，第一个生成步骤留给人类。"Muse"智能体提供专用界面，供设计师输入自己的初始创意。随后，这些人类生成的创意会被转换为与人工智能智能体（"Forge"）所用格式相同的结构化 AOC（行动、对象、情境）格式。

这些人类生成的创意会与人工智能生成的创意一同存入"创意库"。后续的评估智能体"Gatekeeper"对创意来源一无所知。它使用相同且公正的计算指标，对所有创意——无论来自人类还是人工智能——的新颖性和多样性进行评估。设计师可以在任何阶段介入并提供意见，重新运行整个智能体流程。设计师也有权添加人工智能可能已舍弃的创意，和／或移除人工智能已列入候选名单的创意。我们认为，这一机制具有深远的心理影响：

(1) **参与式：** 设计师是创意池的直接贡献者和推动者，而不只是提供提示的人。(2) **主动式：** 设计师的创意不会被搁置；它们会积极地与人工智能的创意竞争，并与之共同演进。(3) **协作式：** 这营造了共同所有权和真正共同创造的氛围。最终筛选出的概念，是不同来源的最佳创意融合而成的混合成果。

这种"PAC"模式确保设计师始终积极参与并处于流程的核心，同时尊重其创作自主权，从而促进更健康、更高效的人机伙伴关系。

3.3 智能体集群：角色与职责

职责

MIDAS 将构思过程分解为多个专业化智能体。每个智能体都由一个大语言模型专门提示，负责一项独立而明确的任务。这种专业化方式类似于人类设计团队，能够实现更稳健、更成熟的工作流程。

3.3.1 阶段*1*：问题定义。 **1. Scribe（问题陈述结构化器）**

- **目的：** 将设计师通常"模糊"且定性的定性问题陈述转化为结构化、计算机可读格式。清晰定义的问题是成功设计的基础 [15]。
- **输入：** 用户提供的非结构化文本问题陈述。
- **过程：** Scribe 解析文本，以识别并提取关键设计参数。
- **输出：** 问题按照我们新颖的 AI3C（活动、项目、矛盾、标准、约束）格式进行结构化处理 [17]。这些结构化数据会保存到"问题库"，供所有其他智能体访问。
- **示例：输入：** "老年人发现从椅子上坐下和站起越来越困难……他们依赖他人……"

输出（AI3C 格式）：

– 活动： 协助完成坐姿与站姿之间的转换。 – 项目： 使用椅子的老年人。– 约束： 必须安全、易于使用、价格实惠，并且适用于家庭环境。–矛盾： 用户需要支持才能独立坐下或站起，但传统椅子无法提供辅助。– 标准：实现独立性、减少外部辅助，并保持舒适性。

3.3.2 阶段 *2*：参与式生成。**2. Muse（创意结构化器）**

- **目的：** 作为人类设计师贡献创意的主要界面，实现"PAC"模型。
- **输入：** 设计师对创意的非结构化文本描述。
- **流程：** Muse 获取设计师的创意，并参考"Scribe"中的 AI3C问题情境，将其重新整理为标准化的创意格式。
- **输出：** 人类创意被整理为 AOC（行动、对象、情境）格式 [17]，并存入"创意库"。
- **示例：输入：** "附着于人体、用于减轻膝部肌肉负荷的仿生结构……"

输出（AOC格式）：

– 创意： 仿生支撑外骨骼– 行动： 连接并提供支撑– 对象：外骨骼– 情境： 减少老年人膝伸肌的肌肉负荷……

3. Forge（创意生成器）

- **目的：** 作为主要的人工智能头脑风暴引擎，模拟人类头脑风暴小组的认知多样性。 • **输入：** 结构化的 AI3C 问题陈述。
- **流程：** Forge 是一个双智能体子系统：– **Formulator：** 一个基于事实的大语言模型（采用较低温度设置），用于生成实用、可行且循序渐进的创意。– **Explorer：** 一个"天马行空"的大语言模型（采用较高温度设置），用于生成突破常规、非传统且高风险高回报的创意。
- **输出：** 一组 10 个新创意（每个子智能体生成 5 个），采用 AOC 格式，并与来自"Muse"的人类创意一同添加到"创意库"中。
- **示例：**

Explorer 输出：
– 创意： 气球云朵椅– 行动： 有节奏地充气和放气，轻轻地将人托起或放下……– 对象： 由响应式气球组成的椅子形簇– 情境： 老年人可以像被云朵托起一般起身或坐下……

Formulator 输出：
– 创意： 可拆卸式助力坐垫– 行动： 在需要时提供额外的支撑力和回弹力。– 对象： 座垫 – 情境： 老年用户放置或移除支撑坐垫……

3.3.3 第三阶段：内部评估和外部基准测试。**4. Gatekeeper（创意评估器 - 局部新颖性与多样性）**

•**目的：** 执行“持续评估”（持续评估）流水线的第一阶段，确保“创意丰沛度”不会变成“创意过载”。评估局部新颖性与多样性。

• **输入：** ‘创意库’中来自 ‘Muse‘ 和 ‘Forge‘ 的所有未经评估的 AOC 创意。

• **流程：** Gatekeeper 使用早期研究 [18] 中提出的语义表征框架，将所有创意转换为向量嵌入，计算它们的余弦相似度，并执行聚类（UMAP + DBSCAN），以识别并筛选出语义上最独特、最多样化的创意。

• **输出：** 经过筛选的高质量 AOC 创意候选名单。

5. Librarian（解决方案研究）

• **目的：** 通过收集现有现实世界解决方案的数据，使构思过程立足于实际，避免系统“重复造轮子”。• **输入：** AI3C 问题陈述。此外，还提供手动模式，允许设计师输入特定链接或已知产品。

• **流程：** 使用基于网络搜索的 LLM，扫描专利、学术论文和商业产品数据库，以查找相关的现有解决方案。

• **输出：** 一份现有解决方案清单，采用 AOC 格式编排（便于直接比较），并附有来源链接以确保可信度，保存至‘文献库’。

• **示例：**
输出（AOC格式）：
– 创意： SitnStand 便携式智能升降座椅– 行动： 通过简单的控制充气和放气，轻柔地抬升……–对象： 电池供电的充气坐垫– 情境： 适用于居家和户外场景，供老年人使用……– 来源： https://www.sitnstand.com

6. Challenger（创意评估器 - 全球新颖性与多样性）

• **目的：** 执行最关键的评估：检查全局新颖性。只有与现有事物不同，创意才真正具有新颖性。

• **输入：** 来自‘Gatekeeper’（‘创意库’）的候选创意，以及来自‘Librarian’（‘文献库’）的现有创意。 • **流程：** 进行语义交叉比较，计算每个生成的创意与全部现有解决方案语料库之间的相似度。任何在语义上与现有产品过于接近的创意都会被筛除。

• **输出：** 一组兼具全球和本地新颖性且多样化的 AOC 创意。

3.3.4 阶段四：发散与可行性。**7.Mint（行动-对象创意生成器）**

• **目的：** 促进“发散”阶段，并打破设计固着。通过将解决方案拆解为核心组成部分，将“从创意中产生创意”这一理念付诸实践。

• **输入：** 来自“Challenger”的全局和局部新颖 AOC 创意。

•**流程：** 解析 AOC 创意，并提取其基本构成要素：一份“行动”列表和一份独立的“对象”列表。• **输出：** 两份相互独立的列表（例如，20 个行动和 20 个对象），作为新的创意“种子”。

• **示例：行动：** 1. 抬升并稳定…… 2. 引导并鼓励…… 3. 支撑并缓冲…… **对象：** 1. 可调节扶手…… 2. 符合人体工程学的过渡垫…… 3. 柔性底座平台……

8. Scout（可行性检查）

• **目的：** 第二个“持续评估”步骤，充当“组合过滤器”，用于管理由“Mint”带来的创意爆发。
• **输入：** 来自“Mint”的动作与对象独立列表，以及 AI3C 问题情境。

• **流程：** 首先，生成所有可能的“动作-对象”对（例如，20 x 20 = 400对）。其次，使用大语言模型为每一对分配“可行性评分”（例如，1/10），评估该特定组合在多大程度上能够解决问题情境。

• **输出：** 按得分排序的列表，其中列出得分最高、最可行的动作-对象对。

• **示例：**
输出：
– 1. 根据用户体重进行变形和调整……- 防滑纹理表面 — 评分：10/10– 2. 通过伸展和回缩提供可调节支撑……- 加固椅背 — 评分：10/10

3.3.5 阶段 *5*：渐进式优化与筛选。**9.Navigator（创意生成器）**

• **目的：** 第二个更加精细的生成阶段，用于实现“从创意中生成创意”。
• **输入：** 来自“Scout”和AI3C问题情境的排名最高且可行的动作-对象对。

• **处理：** 将抽象的动作-对象对“重新还原”为AOC格式中完整且具描述性的创意。此步骤通过综合先前创意中最优秀的部分，生成新的解决方案。随后，它会执行自身内部的新颖性/多样性检查。 • **输出：** 一组全新且经过高度完善的候选AOC创意。

10. Sentinel（相关性专家）

• **目的：** 最终的“筛选”阶段。该阶段用于进行质量保证检查，确保完善后的创意始终与原始问题陈述完全一致。

•**输入：** 来自“Navigator”和“Challenger”的完善后创意，以及来自“Scribe”的原始问题陈述。此外，设计师也可以在此阶段自由选择并输入创意。

• **流程：** 严格执行相关性检查，将每个创意的“情境”与问题的“标准”和“约束”进行比较。润色措辞，并筛除任何偏离原始目标过远的创意。

• **输出：** 最终经过验证、完善且相关的 AOC 创意列表。

3.3.6 第*6*阶段：概念化与可视化。**11. Director（概念生成器）**

• **目的：** 将经过筛选的抽象“创意”转化为具体且可直接实施的“概念”。
• **输入：** 来自“Sentinel”的最终 AOC 创意，以及 AI3C问题情境。

• **流程：** 将简单的 AOC 结构扩展为丰富、详细的**PFIC（原理、特征、实施、特性）**格式 [17]。

• **输出：** 一组结构化的 PFIC 概念，并保存至“概念库”。

• **示例：输入创意：** 带可伸缩辅助装置的集成式安全带**输出（PFIC 格式）：**

> – 原理： 通过可伸缩安全带提供可调节的机械支撑……– 特征： 可伸缩安全带、可调节的张力控制、自动参与……– 实施： 加入电动回缩机构……使用传感器检测用户的动作……– 特征： 安全可靠的支撑、用户友好的自动操作……

12. Leo（图像生成器）

• **用途：** 最后的“概念化”步骤，为这些概念提供逼真的可视化图像。 • **输入：** 来自“Director”的 PFIC 结构化概念。

• **流程：** 将 PFIC 格式中的“特征”和“特征”用作文本转图像生成模型的详细提示。

• **输出：** 一组以视觉方式呈现最终概念的逼真产品渲染图。

3.4 技术实现

MIDAS框架不仅是一个概念模型，还是一个基于“复合大语言模型”方法构建的可运行技术系统，详见图8。

3.4.1 复合式大语言模型方法。 MIDAS 不依赖单一的通用大语言模型，而是采用模型集成，为不同任务选择“合适的工具”。这种智能体方法能够对创作过程进行精细控制。如图8所示，不同的智能体使用不同的模型和参数设置：

• **生成与推理任务**（例如“Scribe”、“Forge Formulator”、“Director”、“Navigator”）可使用强大且连贯的模型，如“Grok 4 Reasoning”，并采用较低的温度（例如“temp=0.5”或“0.6”）以确保精确性。• **创意探索** （例如“Forge Explorer”）使用同一模型，但采用较高的温度（例如“temp=1.0”），以鼓励生成“天马行空”且多样化的输出。• **搜索与事实 grounding**（例如“Librarian”）使用集成网络搜索能力的模型（例如温度为“temp=0.2”的“Sonar Pro”）来获取现实世界的数据。• **图像生成** （例如“Leo”）使用专用的文本转图像模型（例如“gpt-image-1”）。

这种复合方法提供了单个大语言模型无法实现的控制力和专业化程度。

3.4.2 语义评估嵌入。作为「Gatekeeper」「Challenger」和「Scout」智能体重要组成部分的「持续评估」（Continuous Assessment）流水线，由多视图嵌入系统驱动。该系统将文本甚至视觉概念转换为高维向量空间中的表示，从而能够对其创意特质进行数学计算。

• **创意表示：** 「Forge」和「Muse」生成的文本 AOC 输出会被转换为文本嵌入。
• **语义相似度：** 我们主要使用余弦相似度来衡量两个创意向量之间的“距离”。距离越大，意味着新颖性越高。

• **聚类与多样性：** 为衡量多样性，我们在整个创意池上运行聚类算法（使用 UMAP 降维，使用 DBSCAN 进行簇识别）。包含许多规模较小、彼此不同的簇的创意池被视为更多样化。这构成了我们「新颖性、多样性、可行性、相关性」（NDFR）框架的定量基础，该框架基于 [18] 中确立的语义表征方法。

这种计算式评估方法使“CG/CA”循环成为可能，从而让系统能够不断优化自身输出，最终形成真正新颖且多样化的概念集合。

3.5 评估框架：新颖性、多样性、可行性、相关性标准

“更好的创意”这一说法主观且含糊不清。为增强分析的严谨性，我们提出，创意生成系统的成功无法用单一指标（例如“创造力”或“数量”）来衡量。因此，我们在 MIDAS 中整合了一种多维评估启发式方法：**NDFR框架**。该框架

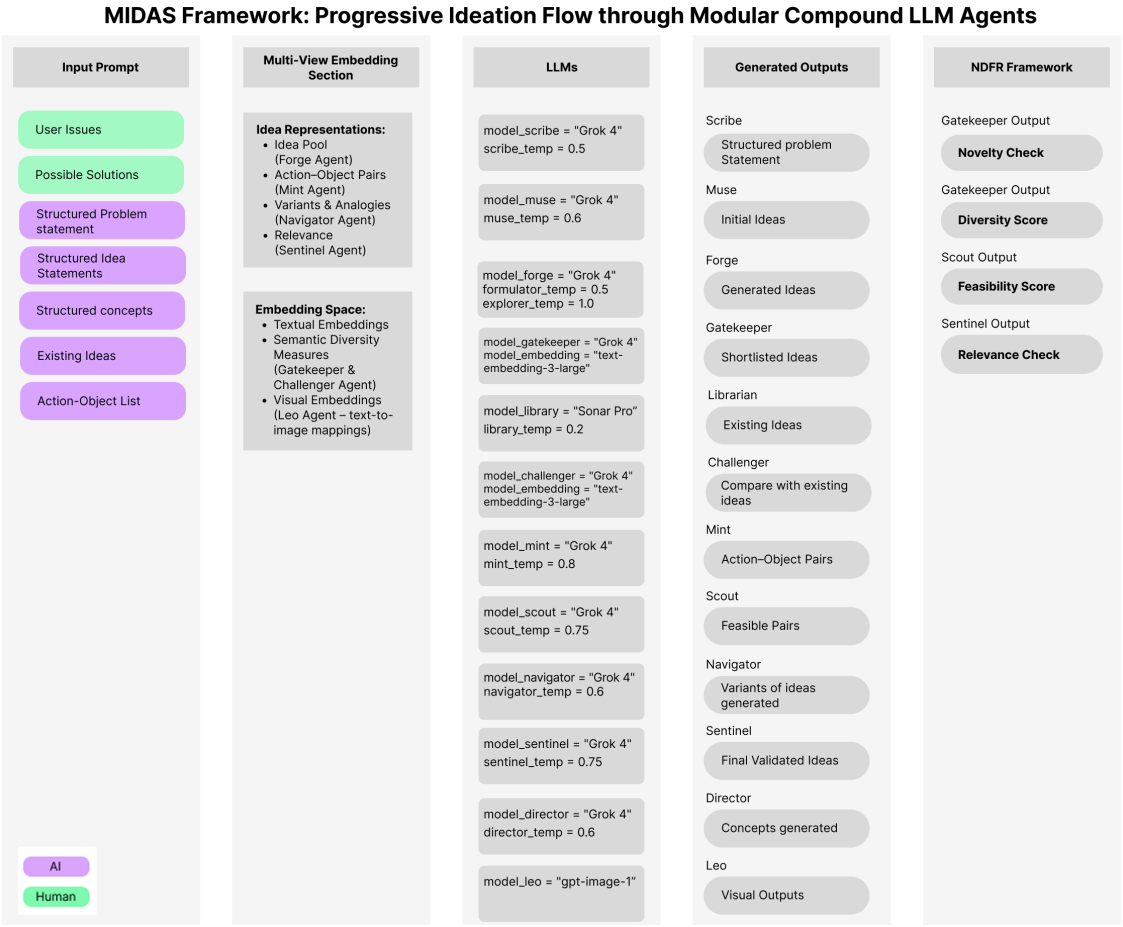


图8：MIDAS框架的技术实现流程图，展示了输入提示、多视图嵌入、分配给各智能体的具体大语言模型、生成的输出以及NDFR评估框架。

从四项标准评估创意：新颖性、多样性、可行性和相关性。

3.5.1 新颖性。我们认为，设计构思中的“新颖性”必须划分为两种截然不同的类型：

- **局部新颖性：**衡量某个创意相对于同一场次中生成的其他创意的原创程度。如果一个系统生成了1000个创意，而这些创意彼此之间都只是细微变化，那么其局部新颖性就很低。该指标由“Gatekeeper”智能体通过计算进行评估，用以量化系统避免自指性固着和语义变体的能力。

- **全局新颖性：**这是衡量创新的真正指标
  它评估一项创意相对于现有的现实世界解决方案整体语料库（例如专利、学术论文和商业产品）的原创性。一项创意可能在局部具有新颖性，但在全局范围内却平淡无奇（即“重新发明轮子”）。这一标准由“Challenger”智能体通过计算进行评估，为创意提供现实依据，并筛选出真正达到最先进水平的创新。

3.5.2 多样性。与新颖性（可适用于单项创意）不同，多样性是整个创意空间的一项属性。它衡量生成的解决方案在语义上的“分布范围”或广度。多样性高的输出会探索多个彼此迥异的概念族群。例如，在解决如何为老年人设计椅子这一问题时，多样性高的输出可能包括机械解决方案、生物识别解决方案、充气式解决方案和材料科学解决方案。多样性低的输出则可能只提供 50 种机械弹簧的变体。我们采用先前研究中确立的基于向量的聚类方法来量化这一指标 [18], ，测量“创意稀疏度”和“簇稀疏度”等指标。

3.5.3 可行性。 一个新颖且多元的创意，如果纯属幻想，就会变成毫无用处的概念。可行性评估的是创意转化为概念时的实践可行性和技术可行性。它要回答：能否利用现有技术和资源实现这一构想？它是否符合问题的基本约束（例如物理学、成本和人体工程学）？ 这一标准由“Scout”智能体明确负责（评估动作—对象配对），也是“PFIC”结构的核心组成部分

由“Director”智能体生成，具体体现在“实施”和“特征”字段中。

3.5.4 相关性。 最后，创意可能新颖、多样且可行，却未能解决实际问题。相关性用于衡量所提概念与原始问题陈述之间的一致程度。它评估解决方案在多大程度上直接回应结构化「AI3C」问题格式中定义的核心「矛盾」「标准」和「约束」。这一目标对齐检查是「Sentinel」智能体的主要功能，可确保渐进式流水线不会偏离用户的核心需求。

这四项嵌入 MIDAS 系统的 NDFR标准共同构成了一个全面、严谨且可通过计算评估的框架，使评估能够超越主观判断，并对生成器智能体所生成创意的产出进行实证评估。

4 研究问题与方法论

如第 3 节所述，MIDAS 的开发基于一个核心假设：分布式、智能体化且渐进式的构思方法优于当前主流的单步整体式大语言模型范式。

4.1 研究问题（R.Q.）

我们的主要研究围绕一个总体研究问题（R.Q.）展开。

R.Q.：基于智能体人工智能的渐进式构思系统（MIDAS）能否生成一组创意，并且有明确证据表明，相较于单阶段、非智能体的人工智能构思方法，这些创意具有更高的新颖性、多样性和可行性?

为全面回答这一问题，我们将主要问题拆解为若干具体的探究方向，这些方向都需要通过我们的方法论加以解决：

（1）**流水线的有效性：**MIDAS中的渐进式多阶段“CG/CA”（持续生成／持续评估）流水线，能否使最终概念集的质量（依据我们的NDFR标准衡量）显著高于单阶段通用型大语言模型交互所产生的“创意丰沛度”?

（2）**新颖性质量：**集成的全局新颖性检查（通过“Librarian”和“Challenger”智能体实现）能否有效筛选出“真正的”创新，生成相对于现实世界解决方案具有新颖性的创意，而非仅仅是语义变体?

(3) **以人为中心的协作：** “PAC”（参与式、主动式、协作式）模型通过“Muse”智能体将人定位为主动的创意贡献者。这种方式如何影响设计师感知到的参与度、所有权感，以及整体共同创作体验的流畅度?

要回答这些问题，需要进行如下所述的充分演示。

5 通过多领域验证展示渐进式构思

为验证 MIDAS 的可扩展性和一致性，我们针对六个不同的设计问题（PS1 至 PS6）开展了一系列全面演示。核心目标是确定智能体渐进式流程能否持续产出经过筛选、真正新颖且多样化的创意，在保持高水平人机协作协同效应的同时，满足严格的 新颖性、多样性、可行性、相关性标准。

在本次展示中，六名新手设计师（设计学硕士生）使用MIDAS平台应对六个问题（PS1至PS6）。与此前的单轮创意丰产实验不同，在那些实验中，设计师花费数小时梳理成簇的相似创意；本研究中的每位设计师每次使用MIDAS的时间最长为*20*分钟。遵循PAC模型，整个过程始于设计师通过*Scribe*智能体界定问题情境。至关重要的是，设计师并非被动旁观者；相反，他们主动贡献了少量初步想法，为*Muse* 智能体启动构思流程，这些想法随后被无缝整合进计算流程。

如表1所详述，MIDAS采用了严格的“筛选与精炼”架构，该表展示了针对六个问题陈述，每个问题陈述中创意通过各智能体的数量。Forge智能体最初生成了大量原始创意（每次会话50至80条不等），随后Gatekeeper根据局部新颖性对其进行筛减。Librarian和Challenger智能体的后续干预确保只有相对于现有文献和专利展现出真正全局新颖性的概念才得以继续推进。最终精选的创意集合（Sentinel输出）十分聚焦，其中PS1（24条）、PS2（19条）、PS3（19条）、PS4（15条）、PS5（11条）和PS6（17条）最为突出。在每种情况下，人工智能生成的贡献在数量上都高于人类提供的初始创意，而最终概念通常被认为是协作式人机对话的自然演进。此外，基于Forge生成的创意，流水线中的后续智能体，即Mint、Scout和Navigator，能够从早先提出的创意中获取新的创意。从表1可以清楚看出，在最终精选的创意总数中，源于解决问题过程中生成的创意与源于创意自身生成的创意几乎均等分布。我们提出的“从创意生成创意”方法在MIDAS框架中取得了极佳效果，证明了不同问题陈述之间存在创意交叉融合的机会。

这些经筛选输出的语义质量通过基于向量的表征框架 [18] 得到了客观验证。如图 9 所示，针对六个问题陈述所生成的创意形成了聚类散点图，其中DBSCAN算法将全部输出归类为单一的“噪声”簇。这一现象在设计研究中具有重要意义，因为它意味着最终集合中的每一个创意在语义上都具有足够的独特性，任意两个创意都无法被归入同一组。这为极高的局部新颖性和多样性提供了实证证据，回应了关于“CG/CA”有效性的首要研究问题

表1：MIDAS智能体集成中的想法筛选与生成数量定量指标。

智能体 / 阶段	PS1	PS2	PS3	PS4	PS5	PS6
Scribe（问题定义）	1	1	1	1	1	1
Muse（人类想法）	8	6	6	5	4	6
Forge（原始人工智能想法）	75	70	65	60	55	78
Gatekeeper（候选名单）	32	28	26	22	18	31
Librarian（现有技术）	10	10	10	10	10	10
Challenger（全球新颖）	11	9	8	7	5	8
Mint（A-O 对）	20	20	20	20	20	20
Scout（可行性评估）	400	400	400	400	400	400
Navigator（新颖／可行）	13	10	11	8	6	9
Sentinel（最终精选）	**24**	**19**	**19**	**15**	**11**	**17**
Director（概念）	22	17	18	14	10	15
Leo（效果图）	20	15	18	12	9	15

流水线。相比之下，先前一次性生成100个创意的研究产生了密集的变体簇（图4、5、6），而MIDAS则交付了少量创意：它们不仅彼此不同，而且相对于由*Librarian* 智能体识别出的现有解决方案而言也具有新颖性。

此外，协作的以人为中心的特性，还体现在*Sentinel* 和 *Director* 阶段的灵活性上。在最终概念化之前，设计师经常会运用自主权进一步完善、补充或删减创意，因此*Director* 和 *Leo* 的最终数量略有差异（表1）。这种主动干预彰显了“持续评估”模型的成功：人工智能并非只是“倾倒”出一个结果，而是真正充当思想启发系统，让人类设计师能够在短短20分钟内引导其朝最有希望且相关的方向发展。

由*Leo*智能体管理的MIDAS流水线的最后阶段，是连接抽象语义表征与具象化概念设计的关键桥梁。系统将高保真度的PFIC（原理、特征、实施、特性）数据转化为照片级真实感效果图，如图10和图11所示，从而将思想启发系统（TPS）理念付诸实践。这些可视化图像不仅呈现最终输出，还作为强有力的视觉线索，触发设计师的横向思维和反思过程。

MIDAS 能够在20分钟工作流程内快速提供高质量效果图，使设计师几乎能够即时评估多种互不重叠的创意在审美与功能上的可行性。这种快速的视觉反馈循环促使设计师以更清晰的思路重新审视流水线的前期阶段，确保人在真正的参与式、主动协作框架中始终处于核心位置并发挥主动作用。

6 讨论

第5节呈现的结果为重新评估人工智能在创意设计过程中的作用提供了有力的实证依据。通过从单体式、一次性生成范式转向分布式、基于智能体且渐进式的工作流程，MIDAS框架回应了早期计算创意构思工具长期存在的“创意丰沛度”根深蒂固的局限。本节综合我们的研究发现，以回应核心研究问题，并探讨基于智能体的人机共创在理论和实践层面的影响。

6.1 通过智能体专业化解构“新颖性幻觉”

本研究的一项主要发现，是单次生成的大语言模型输出所呈现的“表面”新颖性，与通过 MIDAS 流水线实现的“经验证”新颖性之间存在鲜明反差。MIDAS框架通过智能体专业化成功解构了新颖性幻觉。通过将构思过程拆分为不同角色，系统得以超越简单的插值。例如，*Forge-Explorer* 智能体经过专门参数化，以绕过高概率词元；而 *Mint*和 *Scout* 智能体则通过组合合成，促成“从创意中产生创意”。图9中观察到的“噪声”簇正是系统最显著的成功之处；它们表明系统成功拓展了概念空间的边界，使得任意两项创意之间都不存在显著的语义重叠。这证实，MIDAS流水线实现的局部新颖性 和多样性 在数学和创造性层面均优于传统的单次突发式生成。

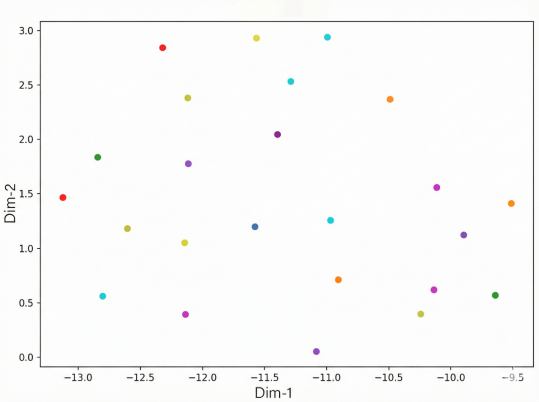
6.2 时间性优势：将渐进式认知付诸实践

我们的研究强调了创造力一个根本却常被忽视的维度：时间性。我们指出，人类构思是一个路径依赖过程，其中的（$n+1$）th 创意，是对其前序 n 创意的反思性演进。现有人工智能系统在单次生成中产出数百个创意，违背了这种时间逻辑，导致“概念停滞”：模型实际上会在单个上下文窗口内重复自身。

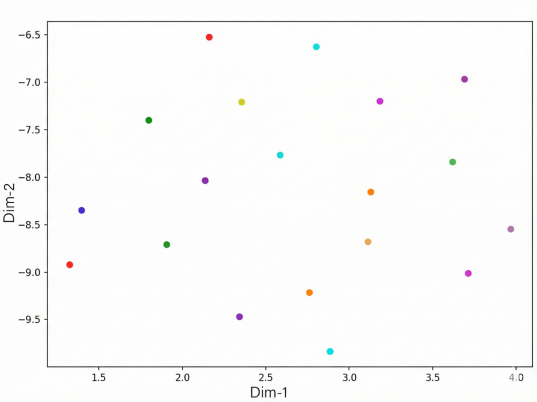
MIDAS框架通过其渐进式流水线将这种时间顺序付诸实践。*Navigator* 智能体并非凭空生成创意，而是在*Mint* 拆解出的动作-对象对以及*Scout*的可行性过滤器基础上继续构建。从模糊的问题陈述到结构化创意，再到拆解后的组件，最终综合形成概念，这段循序渐进的过程模拟了人类设计师的元认知“循环”。针对PS1–PS6的实验结果（图4–6）证实，这种“渐进式激发”所产生的创意，始终比任何一次性迸发所能实现的结果更远离初始创意群。这表明，打破人工智能设计固着的关键不在于使用“更大型”的模型，而在于采用“更缓慢”、结构更清晰且具有时间意识的交互方式。

6.3 重建设计师能动性：对“PAC”模型的反思

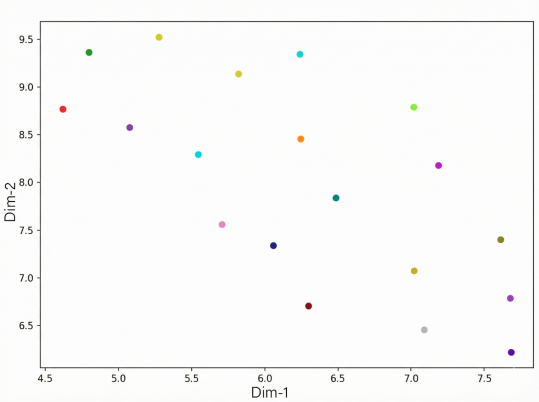
或许，本研究最显著的定性成果，是“*PAC*”（参与式、主动式、协作式）模式的成功。在当前的人工智能辅助设计文献中，专业人士普遍表现出明显的“人工智能抵触”，他们认为生成式工具



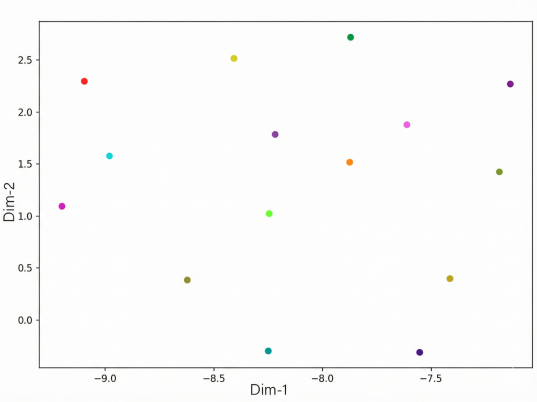
(a) PS1：废弃物管理（24个创意）



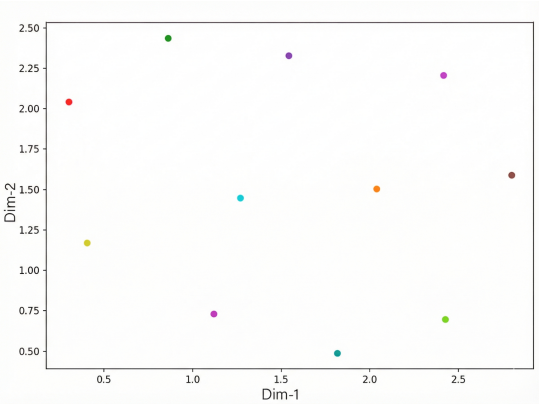
(b) PS2：鞋类消毒（19个创意）



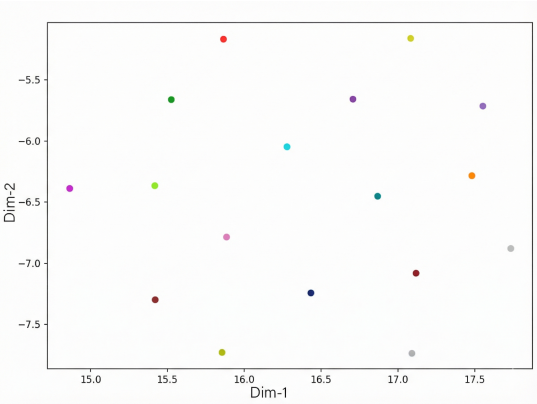
(c) PS3：洗碗（19个创意）



(d) PS4：站立舒适度（15个创意）



(e) PS5：喂鸟（11个创意）



(f) PS6：雨伞收纳（17个创意）

图9：使用 MIDAS框架对六个设计问题中的创意进行聚类分析。对于所有问题，DBSCAN算法都将整个创意池识别为“噪声”，这表明该系统始终能够生成高度多样化、彼此不重叠的概念集合，打破了单体大语言模型输出中常见的语义固着。

贬低了他们的专业能力。将人类降格为机器生成的“垃圾内容”的被动筛选者，这些工具会削弱人们的参与度，并使其丧失创作自主权。

MIDAS框架扭转了这种权力关系。将*Muse* 智能体纳入首个生成步骤，确保设计师的“创意火花”成为整个流程的基础。



(a) PS1：可折叠分类套筒



(b) PS1：天然酶喷洒器



(c) PS1：防漏漏斗屏障



(d) PS2：太阳能紫外线手套



(e) PS2：净化蒸汽室



(f) PS2：光催化除臭器



(g) PS3：生物旋流水回收器



(h) PS3：智能水龙头适配器



(i) PS3：逆流喷嘴

图 10：Leo 代理为 PS1–PS3 生成的概念效果图。这些可视化图像展现了渐进式构思流程的最终成果，其中结构化的 PFIC 概念被转化为逼真的图像，以引发设计师的最终反思。

在我们的案例演示中，设计师表示，他们之所以有很强的所有权感，正是因为自己的创意在“创意库”中被视为与机器贡献同等重要。*Sentinel* 而*Director* 智能体则进一步强化了这一点，使设计师能够筛选并完善最终的概念组合。这种协作流畅性表明，人工智能的力量并非在于



（a）PS4：可折叠迷你凳



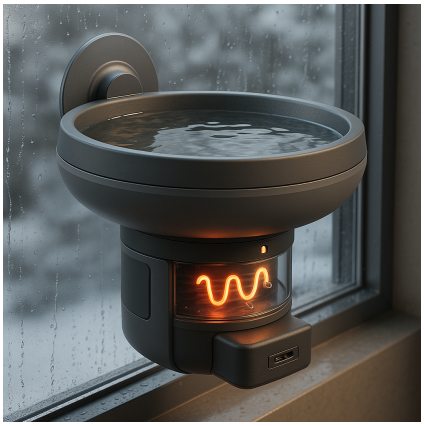
（b）PS4：环保泡沫鞋垫



（c）PS4：自调节姿势底座



（d）PS5：粘贴式窗户喂食器



（e）PS5：冬季温水装置



(f) PS5：防冻水源



(g) PS6：手风琴式挤压包裹带



(h) PS6：离心式拉伸旋转器



(i) PS6：被动热循环锁扣

图 11：Leo 代理为 PS4–PS6 生成的概念效果图。这些可视化图像展现了渐进式构思流程的最终成果：结构化的 PFIC 概念被转化为逼真的形式，以引发设计师的最终反思。

取代人类，而在于充当思想启发系统（*TPS*），成为一种认知支架，引导人类探索尚未涉足的方向

各领域，同时尊重其作为价值与相关性的最终裁决者所发挥的作用。

6.4 弥合差距：从语义流畅性到全球创新

当前人工智能构思研究的一个关键弱点，是缺乏“现实世界的根基”。在聊天窗口这一狭窄情境中具有新颖性的创意，与人类发明的悠久历史相比，可能显得微不足道。MIDAS 通过 *Librarian* 和 *Challenger* 智能体弥合了这一差距。通过将实时网络搜索和专利基准比较整合到流水线中，该系统实现了从“语言创造力”到“技术创新”的转变。

表1中的筛选指标揭示了显著的“漏斗效应”。尽管*Forge*可能会生成 70+ 创意，*Challenger*智能体通常会将其筛选至少数几个具有全球新颖性的候选方案。这项严格的全局新颖性检查确保设计师只将时间花在有望达到真正先进水平的概念上。在“知名会议”的情境中，这种扎根于现有技术的做法至关重要，因为创新是以现有技术为基准衡量的。MIDAS能够在20分钟内持续产出此类高质量、具有全球新颖性的成果，这标志着计算机辅助创新（CAI）在效率和效能方面取得了重大飞跃。

6.5 对设计实践与教育学的启示

MIDAS的成功对设计实践和教育都具有深远影响。对于行业从业者而言，MIDAS提供了一种克服“评估过载”的方法。设计师无需再花费数小时梳理充斥着各种变体的头脑风暴记录，而是可以在不到半小时内获得包含10至15个高质量、可行且新颖概念的多元组合。这种效率使设计过程的早期阶段能够开展更频繁、更深入的迭代。

在教学情境中，智能体的结构化特性为学生提供了一份“认知路线图”。通过与 Scribe（用于问题定义）或 Mint（用于发散思维）等智能体互动，新手设计师在受益于人工智能庞大知识库的同时，也能学习优秀设计的方法论。因此，MIDAS 既是生产工具，也是学习工具，能够强化顶尖设计院校所教授的系统化设计方法 [15]。

6.6 局限

尽管试点研究的初步结果令人鼓舞，但我们必须承认这项初期工作的局限。

- 当前系统依赖基于文本和图像的大语言模型；未来版本可以通过整合能够对三维几何或功能物理仿真进行推理的多模态智能体而获益。
- 此外，尽管 20 分钟的时长是一项优势，但仍需开展纵向研究，以了解持续使用此类系统会如何影响设计师长期的创造性习惯。
- 本研究针对一个规模较小且构成单一的群体开展，该群体由六名新手设计师组成（n=6）。虽然这提供了丰富的定性反馈，但研究结果不能在未经进一步验证的情况下推广至更广泛的资深设计师或行业专业人士群体。
- 这些设计问题虽然复杂，但都是界定明确的产品设计情景；该框架在更模糊、系统性更强或“棘手”的问题上的表现仍有待检验。
- 混杂变量，例如 MIDAS 系统本身的新颖性，或研究的顺序设计所产生的潜在学习效应，都值得通过控制更严格、规模更大的研究进一步考察。

尽管如此，六个问题的研究结果都为智能体方法提供了有力验证。我们证明，将构思视为渐进式、多智能体、参与式的旅程，能够让我们超越当前人工智能的“随机中位数”，开启真正创新的人机共创新时代。

6.7 未来研究方向

本研究基于研究中的局限与见解，为未来探索开辟了多个前景广阔的方向。

(1) **大规模验证：**下一步应立即开展规模更大、更多样化的验证研究。这包括招募专业水平各异（例如，新手学生与拥有20年行业经验的资深人士）、来自不同领域（例如，工程、服务设计、用户体验/用户界面）的参与者，以评估该框架更广泛的效用和适应性。(2) **改进智能体能力：**未来的工作将着重增强各个智能体的专业化程度。例如，可以增强“Librarian”智能体，使其能够解析包含CAD文件和技术图的多模态专利数据库，而不仅仅是文本。“Scout”智能体则可以接入工程知识图谱或基础物理仿真，从而提供更稳健的可行性评分，超越基于大语言模型直觉的判断。(3) **深入研究人机交互：**“PAC”模型值得开展专门的人机交互研究。纵向研究可以探索设计师持续使用MIDAS系统后，其创作过程和认知技能如何随时间演变。我们还希望探索自适应能动性，即系统可根据检测到的设计师专业水平或认知状态，动态调整“激发”或支持的程度。(4)**混合智能架构：**虽然MIDAS目前采用复合式大语言模型方法，但我们将探索为特定智能体整合非大语言模型人工智能技术，例如符号推理引擎或进化算法。混合智能“Scout”（用于可行性评估）或“Sentinel”（用于相关性评估）能够将生成流畅性与逻辑严谨性有力结合。(5) **从可视化到制造：**当前流水线的输出止于“Leo”（图像生成）。未来的一项重要目标是扩展这条流水线，开发一个“Fabricator”智能体，将最终的PFIC 概念转化为初步CAD模型、3D可打印文件或物料清单，从而弥合概念与实体原型之间的差距。

总之，MIDAS是迈向新一代智能体化、渐进式且真正协作式设计工具的重要基础。未来的发展不仅需要改进这项技术，也需要加深我们对人与人工智能之间新型创造性伙伴关系的理解。

7 结论

本文探讨了单次突发式、单体式人工智能系统在创意构思这一复杂、多阶段流程中的不足。研究表明，将单一人工智能“伙伴”转变为由专业协作智能体组成的分布式“团队”，是实现人机协同创造并生成真正新颖的一组创意的更有效模式。本文的主要贡献在于提出了 MIDAS 框架及系统。MIDAS 将用于人机伙伴关系的“PAC”模型付诸实践，尊重设计师自主权并促进共同所有权。智能体流程通过“CG/CA”工作流程模拟元认知，逐步将创意完善为创新概念。因此，本研究展示了一条实现协同伙伴关系的可行路径：机器负责复杂分析与广泛探索，使人类能够专注于提供情境、行使判断并引导创意历程。综上所述，MIDAS 代表了面向人机共创的新一代智能体化、渐进式、主动式、参与式且真正协作式的设计范式。

参考文献

[1] Luca Abrusci、Karma Dabaghi、Stefano D’Urso 和 Filippo Sciarrone。2025年。AI4Design：一种基于生成式人工智能的系统，用于提升设计中的创造力——一项实地评估。《计算机与教育：人工智能》8（2025），100401。doi: 10 1016 2025 100401 。/j.caeai。。Emre Can Acikgoz、Jinoh Oh、Jie Hao、Joo Hyuk Jeon、Heng Ji、Dilek Hakkani-Tür、Gokhan Tur、Xiang Li、Chengyuan Ma 和 Xing Fan。2025年。SpeakRL：通过强化学习协同语言模型中的推理、表达与行动。arXiv: 2512 13159 [ ] 2512 13159 。cs.AI https://arxiv.org/abs/。Ammar Ahmed 和 Ali Shariq Imran。。大语言模型在 UI/UX设计中的作用：系统性文献综述。*arXiv*预印本 *arXiv:2507.04469* （2025）。[4] Nathalie Bonnardel 和 John Didier。2020年。促进创意设计的头脑风暴变体。《应用人体工程学》83（2020），102987。doi:10.1016/j.apergo.2019。102987[5] Filippo Chiarello、Simone Barandoni、Marija Majda Škec 和 Gualtiero Fantoni。2024年。工程设计中的生成式大语言模型：机遇与挑战。《设计学会论文集》4，1959–1968。doi:10.1017/pds。2024.198[6] Mohammad Mohiuddin Choudhury、Boris Eisenbart 和 Blair Kuys。2025年。设计过程中的人工智能（AI）——关于生成式人工智能视角的综述与分析。《设计学会论文集》5，631–640。doi: 10 1017 2025 10077 。/pds。。Julian De Freitas、Gideon Nave 和 Stefano Puntoni。2025年。借助生成式人工智能进行构思——在消费者研究及其他领域。《消费者研究期刊》52，1（2025），18–31。[8] Zhoumingju Jiang、Ang Liu、Dawen Zhang、Xiwei Xu 和 Yun Dai。2025年。面向工程设计的大语言模型定制化与个性化。《*CIRP*年鉴》 74, 1 2025 191 195 10 1016 2025 03 001 （），–。doi：。/j.cirp。。C.K.M. Lee、Jingying Liang、K.L. Yung 和 K.L. Keung。。利用生成式人工智能生成受TRIZ启发的生态设计指南。《高级工程信息学》 62 2024 102846 10 1016 2024 102846 （），。doi：。/j.aei。。Yiren Liu、Pranav Sharma、Mehul Oswal、Haijun Xia 和 Yun Huang。。PersonaFlow：设计大语言模型模拟的专家视角，以增强研究构思。（2025 506 534 10 1145 3715336 3735789 ），–。doi：。/。Yiren Liu、Pranav Sharma、Mehul Oswal、Haijun Xia 和 Yun Huang。2025年。PersonaFlow：设计大语言模型模拟的专家视角，以增强研究构思。载于*ACM*交互系统设计会议（*DIS ’ 25* 10 1145 3715336 3735789 ）。葡萄牙丰沙尔。doi：。/。Kevin Ma、Daniele Grandi、Christopher McComb 和 Kosa Goucher-Lambert。2024年。大语言模型能生成多样化的设计概念吗？与人工众包解决方案的比较研究。《计算学期刊》 与工程信息科学25，2（2024年12月），024501。doi:10.1115/1。4067332[13] San S. Nguyen 和 Javeed Kittur。2025年。大语言模型在工程领域当前及潜在用途的综述。*Frontiers in Education* 第10卷—2025年（ 2025 10 3389 2025 1649650 ）。doi：。/feduc。。Christian Obieke、John Bridgeman 和 Ji Han。2025年。工程设计中的人工智能协作框架（AICED）。《设计学会论文集》5，91 100 10 1017 2025 10023–。doi：。/pds。。国际工程设计会议，ICED25；会议日期：2025年8月11日至2025年8月14日。[15] Gerhard Pahl 和 Wolfgang Beitz。1996年。工程设计：系统方法（第3版英文版）。Springer-Verlag London，伦敦。[16] Christian Pescher 和 Gerard J. Tellis。2025年。人工智能在构思过程中的作用。产品创新管理期刊42，5（2025），958 982 10 1111 12791 –。doi：。/jpim。B Sankar 和 Dibakar Sen。2025年。利用结构化对话式人工智能（CAI）系统实现的新型创意生成工具。人工智能与工程设计、分析及制造39（2025），e11。doi:10。1017/S089006042500006X[18] B. Sankar 和 Dibakar Sen。2025年。构思的语义表征：一种使用向量嵌入和聚类的计算框架。工程设计期刊 0, 0 2025 1 45 10 1080 09544828 2025 2537465 （），–。doi：。/。。Aswathy Sreenivasan 和 M. Suresh。。设计思维与人工智能：探索二者协同作用的系统性文献综述。国际创新研究期刊 8, 3 2024 297 312 10 1016 2024 05 001 （），–。doi：。/j.ijis。。Robert I. Sutton 和 Andrew Hargadon。。情境中的头脑风暴小组：产品设计公司的有效性。《行政科学季刊》41，4（1996 685 718 2393872 ），–。http://www.jstor.org/stable/Hsien-Hui Tang、Ying-Ling Chen 和 John Gero。。设计方法对设计过程的影响：情景、头脑风暴和综摄法在设计中的使用效果。（07 2012 10 21606 2012 134 ）。doi：。/drs。。John Terninko、Alla Zusman 和 Boris Zlotin。。系统创新：*TRIZ*（发明问题解决理论）导论。CRC出版社。[23] Y.P. Tsang 和 C.K.M. Lee。2022年。工业设计中的人工智能：半自动化文献调查。人工智能工程应用 112 2022 104884 10 1016 2022 104884 （），。doi：。/j.engappai。。Samangi Wadinambiarachchi、Ryan M. Kelly、Saumya Pareek、Qiushi Zhou 和 Eduardo Velloso。2024年。生成式人工智能对设计固着和发散思维的影响。载于*2024*年计算机系统中的人为因素*CHI*会议论文集（美国夏威夷檀香山）（*CHI ’24*）。计算机协会，美国纽约州纽约市，文章380，18页。doi:10.1145/3613904。3642919[25] Nan Wang、Hyunsuk Kim、Junfeng Peng 和 Jiayi Wang。2025年。探索人机共创中的创造力：跨设计经验的比较研究。*Frontiersin ComputerScience* 第7卷—2025年（2025）。doi:10.3389/fcomp.2025.1672735[26] Pan Wang、Yash Khinvasara、Geesje Josine Creijghton、Tessa Scholing、Yihua Wang、Zhibin Zhou、Peter R.N. Childs 和 Yuan Yin。2025年。通过人机协同构思提升设计师创造力：利用定制GPT进行设计构思的共创框架。工程设计、分析与制造人工智能39（2025），e22。doi:10.1017/S0890060425100127[27] Pan Wang、Xun Zhang、Liyan Wei、Peter Childs、Stephen Jia Wang、Yike Guo 和 Maaike Kleinsmann。2025年。通过组合生成模型实现人机协同构思。工程设计期刊0，0（2025），1–37。doi:10.1080/09544828。2025.2504309[28] Zuoxu Wang、Xinxin Liang、Mingrui Li、Shufei Li、Jihong Liu 和 Lianyu Zheng。2025年。迈向认知智能赋能的产品设计：人工智能赋能产品设计的演进、最先进水平与未来。工业信息集成期刊43（2025），100759。doi:10.1016/j.jii.2024.100759[29] Shunyu Yao、Jeffrey Zhao、Dian Yu、Nan Du、Izhak Shafran、Karthik Narasimhan 和 Yuan Cao。2023年。ReAct：协同语言模型中的推理与行动。arXiv: 2210 03629 [ ] 2210 03629 。cs.CL https://arxiv.org/abs/。Andy Zhou、Kai Yan、Michal Shlapentokh-Rothman、Haohan Wang 和 Yu-Xiong Wang。2024年。语言智能体树搜索：统一语言模型中的推理、行动与规划。arXiv: 2310 04406 [ ] 2310 04406 。cs.AI https://arxiv.org/abs/。Yehong Zhou 和 Chun-Hsien Chen。。考察大语言模型对设计的影响：功能、优势、局限与角色。设计与人工智能1，2（2025），100017。doi:10.1016/j.daai.2025.100017