arXiv:2305.03852v1 [cs.HC] 5 May 2023

CHAI-DT：一种提示对话式生成式人工智能智能体积极参与共同创造的框架

BRANDON HARWOOD, IBM，美国

摘要： 本文[1] 探讨在以群体为中心的共创框架中运用生成式人工智能模型的潜力，以提升商业创新与共同创造情境中的问题解决能力和创意构思，并提出一种新颖的提示技术，用于对话式生成式人工智能智能体。这种技术借鉴传统“人与人”引导和指导方法，使智能体能够积极参与设计思维这一共创框架。通过运用该提示技术开展实验，我们收集到证据，表明对话式生成型转换器（即 ChatGPT）能够为设计思维活动提供符合情境、实用且富有创意的意见。我们还讨论了在共创构思中使用生成式人工智能模型的潜在益处、局限与风险，并对未来研究提出建议。

ACM 概念： • **以人为中心的计算** → **协作交互**。

其他关键词和短语：人机交互、人工智能、生成式人工智能、共同创造、设计思维、

ACM 参考文献格式：
Brandon Harwood。2023。CHAI-DT：一种通过提示对话式生成式人工智能智能体主动参与共同创造的框架。1，1（2023 年 5 月），7 页。

1 介绍

近年来，人工智能技术的发展及其可及性的提升，使得 ChatGPT 等生成式人工智能模型 [8] 能够作为多种情境和领域中的生产力、研究及创意应用的通用工具 [7]，并成功参与用户体验设计职能。 [5] 尽管这些应用场景通常是在一对一的人类 + 人工智能交互情境下进行研究的，但很少有研究探讨参与者和引导者如何在以群体为导向的共同创作框架中有效利用生成式创意智能体，例如设计思维。本文认为，这类应用能够为商业创新和共同创造情境中的问题解决与创意构思带来价值。 [6]

本文旨在探索群体 + AI 共创伙伴关系的空间，并提出一种提示技术，使对话智能体（即 ChatGPT）能够与人类共同创作。该技术借鉴了传统「人与人」引导和指导中所采用的方法，这些方法通常见于设计思维——一种共创框架，旨在让来自不同学科的参与者在团队中协作并共同构思，为商业领域构建以用户为中心的解决方案。 [3] 通过使用这项提示技术进行实验，我收集到证据，表明像 ChatGPT 这样的对话式生成 Transformer 能够为设计思维活动提供符合具体情境、有用且富有创意的内容。我

[1]这是 本文是在 ACM CHI 2023 生成式人工智能与人机交互研讨会上首次发表的论文基础上，经过少量修订和扩展而成的版本 .[2]

作者地址： Brandon Harwood，IBM，达拉斯，得克萨斯州，美国，brandon.harwood@ibm .com.

允许出于个人或课堂用途免费制作本作品全部或部分内容的数字或纸质副本，前提是这些副本不得用于营利或商业利益，不得进行复制或分发，且副本首页须附上此声明和完整引文。本作品中由 ACM 以外的其他方拥有版权的组件，必须遵守其版权规定。注明出处后允许摘录。以其他方式复制、再版、发布到服务器或分发至列表，均须事先获得明确许可和/或支付费用。请联系 permissions@acm.org 申请许可。
© 2023 美国计算机协会。提交至 ACM 的手稿。

还将探讨该框架在 Kantosalo 的人机共创 5C 框架和 IBM 设计思维框架中促进“人类/ + 人工智能”协同创作伙伴关系的潜在价值，相关内容详见 [3]。 [6]

最后，我将探讨在共创构思中使用生成式人工智能的潜在益处、局限与风险，并展望未来研究的机会，思考如何降低风险，以及衡量“人类／ + 人工智能”共创伙伴关系的成效。通过应对这些局限，并探索在群体环境中与创造性人工智能智能体有效合作的方式，我希望为未来的发展指明方向，增进我们对人工智能系统／智能体的理解，并推动其在共同创造活动中成为有帮助且有效的合作伙伴与参与者。

2 背景

Kantosalo 等人将人机协同创造描述为：“人类与计算机集体内部的互动、该集体的协作过程，以及对社群的创造性贡献，这些都处于丰富的情境之中。”[4] 这一定义与传统的人类共创框架（例如设计思维（DT））形成对比，但也与之密切相关；二者的区别在于，前者明确考虑了计算智能体在共同创造中的贡献。以 Kantosalo 的模型为参照，设计思维可描述为：由一群集体组成的个体，运用各自“主观的”情境，围绕一个社群（例如用户或用户群体，他们通常作为参与者参与设计思维工作坊），在设计思维环节所设定的范围狭窄的“理想化目标”情境中（例如问题空间、流程等），彼此协作，并通过 [3], 中所述的设计思维框架，最终贡献创意成果（例如想法、观察等）。

Kantosalo 指出：“在成功的创造性协作中，集体能够从集体内部深刻的沟通以及贡献分享中获益。” [4] 该框架通过明确“人类主动性与计算主动性”二分法而区别于设计思维，但这一成功标准同样可以用来描述成功的设计思维环节，唯一的区别在于，在传统设计思维中，只有人类参与者拥有创意主动性。然而，近期关于生成式人工智能智能体通过对话为设计空间创造性地作出贡献的能力的研究 [5], 表明，我们能够构建一个兼顾人类主动性与计算主动性的系统，并将创造性人工智能智能体作为积极参与者整合到群体创意构思空间中。

基于上述背景，我希望通过一种新颖的提示框架，探索将创造性人工智能智能体融入设计思维的潜力，并进一步探讨如何通过人机混合主动性，在以群体为中心的共同创造环境中构建能够促进更多创意产出的环境。

3 方法

3.1 初步尝试与发现

由于 OpenAI 的 ChatGPT [8] 易于获取、能力相对先进，且具备对话界面，因此我们选择它作为实验测试平台，为设计思维引导者和参与者在实时构思过程中与模型互动提供了环境。在初步测试中，我们发现，通过松散的对话式互动来“引导”模型完成设计思维练习，会导致输出质量不稳定，需要多次来回解释或重新措辞提示。这种方式足以满足一对一构思的需要（有时甚至是理想选择），但在实时活动中可能会令人分心。

然而，通过这些松散的互动可以清楚地看出，（与人类参与者颇为相似）如果向模型清晰说明活动内容、介绍活动适用的背景、提供分步指令，并说明创意构思为何从一开始就需要进行，模型便能更成功地完成设计思维练习。

这些模式有助于形成两组指令和说明；将其提供给 ChatGPT 后，模型便能生成更出色、结构更清晰的协同创作内容。此外，在与模型互动之初提供这些指令和说明，可以让引导者只需指示模型执行该练习（或练习中的某一步或后续步骤），同时使引导者和人类参与者能够采纳、拒绝或补充模型生成的创意成果，并继续进行创意构思。

3.2 CHAI-DT 框架

我所介绍的**协作式人机设计思维（CHAI-DT）提示方法**是一种便捷且灵活的框架，它将这些指令和解释整合到单个提示中，用于启动互动，并使对话代理能够在设计思维活动中为小组贡献有用的创意成果。具体而言，该方法通过说明活动、定义成功标准，并通过六个子提示的组合提供具体指令和必要情境来实现这一点。前四个子提示是“静态指令”，以我们向人类参与者介绍活动的相同方式，介绍所选活动的结构和目的，并且不会因与代理的互动而改变。最后两个是“可变指令”，取决于引导者和参与者的会话情境及互动需求。

静态指令：

- 介绍：“我们正在开展一项‘[活动名称]’设计思维练习。”
- 定义：解释该练习、其预期成果，以及其为何有助于共同创造。
- 示例：一个或多个理想最终成果的示例。
- 指令：活动的分步指令。

可变指令：

- 情境：为互动提供适当的背景。（即：我们为什么要共同创作？谁会参与？我们可以提供哪些背景信息？）
- 执行：提示模型完成活动，可以一次性完成，也可以逐步完成（例如：“完成整个 HILLS 练习。”或“完成第 1 步。”）。

4 结果

4.1 设计思维“Hills”

下方的表1记录了 ChatGPT 2 月 13 日发布版 [9], 在 CHAI-DT 框架提示下，逐步完成“Hills”设计思维练习的输出（提示词可在附录中查看）。

Hills 是“我们旨在帮助用户实现目标的简洁陈述” [3], ，参与者通过观察和反思一个狭窄的情境来构建这些陈述，并各自贡献成果（通常写在便利贴上，每张写一个想法），这些成果需符合以下 3 项标准：

- **“谁”**，即我们所设计的情境中的相关人员。
- **“什么”**，即为“谁”这一部分中的人员提供的赋能功能。
- **“哇”**，即我们的解决方案所带来的差异化价值或影响。

通常，在参与者根据每项标准创作出许多成果后，团队会将各个成果归纳为主题组，并反思由此浮现的主题，以协作构思创意。这些主题组有助于共同创作，因为它们能够整理并凸显团队中多种视角所提出的想法。

表1. ChatGPT 对 3“山丘”设计思维练习的回答 步骤

“谁”（潜在用户）	**“什么”（赋能功能）**	**“哇”（影响）**
零售店经理	准确预测销售趋势	大幅减少缺货和过度 库存，从而提高销售额和 盈利能力
库存经理	实时优化库存水平	通过确保.....来提升客户满意度 产品始终有库存
供应链经理	找出表现不佳的产品	提高效率和生产力，适用于 门店和库存经理
销售人员	识别库存过剩的产品	减少浪费并优化资源利 用
顾客（间接受到影 受系统影响）	确定补货数量和时间 ing	通过先进的数据分析， 在零售行业取得竞争优势 和人工智能
高管/决策者 RetailInc 的决策者	监控库存水平，并在库存低于某一 阈值时提醒管理人员 阈值	确保销售额和库存 并进行盘点 数据，从而获得更好的 决策
	提供有关客户需求的见解。 命令与行为	提升整体购物体验 通过改善库存为客户提供更好的服务
	生成自动化报告和分析- 用于库存和销售数据的分析	管理与产品供应情况 使 RetailInc 能够快速响应无码高清 不断变化的市场趋势和客户需求- mands
	最大限度减少缺货和库存积压	在 RetailInc 培养创新和持续改进的文化 在 RetailInc 培养创新和持续改进的文化
	为库存管理提供数据驱动的决策制定 支持	

一次性指示 ChatGPT 完成此练习时，结果不错。然而，当提示模型逐步完成这项活动时，我们能够测试它反思所提供情境以及小组／其自身先前产出的能力，从而为小组生成观察性产物，例如根据情境推导出的用户，以及创意性产物，例如他们可能面临的问题和对他们有价值的新想法。允许模型反思自身先前的输出后，指示模型逐步完成 Hills活动所得到的整体结果显示，产物的质量更高，也更切合实际。

虽然表1中的输出来自一次单独的一对一测试，但它们代表了多次测试中常见且可重复的模式。根据这些结果，我们可以看到，ChatGPT 能够生成与会话情境相关的内容，在质量、创造性以及对协作式创意构思的实用性方面表现合理。模型的回复也遵循了所提供的指令，并且可以轻松融入实时的设计思维会议中。

4.2 情境

本次测试中提供的情境子提示是由 ChatGPT 1月9日发布版生成的虚构场景 [9], ，仅用于测试不同的设计思维活动：

大型零售连锁企业 *RetailInc* 与 *IBM* 携手改进其库存管理系统。现有系统已经过时，常常导致缺货和库存积压，进而造成销售损失和资源浪费。*IBM* 为此次合作带来了数据分析和人工智能方面的专业知识，而 *RetailInc* 则提供了其大量关于

销售额和库存水平的数据。双方旨在共同打造一套新系统，准确预测销售趋势并实时优化库存水平，从而提高零售连锁企业的效率和利润。

4.3 “自我”认可与协作行为

有趣的是，ChatGPT 完成 Hills 练习的每个步骤后，还特意添加了免责声明：

“注：这些只是一些潜在用户 *[*、‘系统可以提供的赋能功能’或‘系统可以提供的潜在市场价值或差异化因素’*]*，，团队可能需要根据其对用户需求的研究和理解，进一步完善这份列表。”。

这种行为将生成的内容塑造为旨在 与他人协作的内容，并通过促成人类/ +AI 集体 之间的互动，使其符合 Kantosalo 框架，即通过个人贡献 ，围绕所提供情境 中所识别的社群 进行协作 。

5 讨论

通过开发 CHAI-DT 提示框架，我证明了 ChatGPT 有能力在实时共同创造活动中提供实用且富有创意的成果。这种提示工程方法的局限在于，它应用于企业协同创作的情境，而企业协同创作通常用于创造商业价值；不过，该方法也适用于多种不同形式的共同创造式、基于指令的协作，未来还需要进一步研究如何将其融入其他以群体为中心的共同创造活动。然而，这种人类与 + AI 共同创造模式的影响深远，为未来研究人或群体如何与生成式人工智能智能体协作、结成伙伴关系并互动，以提升其他领域和情境中的集体创造力提供了空间，同时也为旨在实现协作、合作与共同创造的人工智能智能体的后续设计探索提供了参考。

通过为符合 Kantosalo 的 5Cs 框架的共同创造型生成式智能体引入现实世界应用，我推断，采用类似或迭代的方法有机会进一步探索这一领域。在共同创造中运用 CHAI-DT 等方法，尤其是与支持同生成式人工智能智能体开展更高级互动的其他系统相结合，将开启全新的创造性协作形式，并创造出有助于加强跨学科沟通与共同创造力的环境或方法。

未来研究还应探讨在这些环境中使用生成式人工智能智能体可能带来的风险。Buschek 等人已对此进行了充分论述，例如偏见、错误信息、责任冲突，以及可能接触私人或不当数据 [1], ；此外还包括参与者分心、创意构思效果欠佳或造成危害、恶意行为者滥用该框架，以及对我们进行创意构思所服务的社群造成潜在伤害。

6 结论

本文介绍了 CHAI-DT，这是一种新颖的方法，旨在通过提示 ChatGPT 等对话智能体参与“人类与人工智能”协同创作；并探讨了将创意型生成式人工智能智能体融入实时设计思维的潜力，以及这一方法的影响和相关潜在风险。我还建议开展更多研究，考量此类框架（无论有益还是有害）的风险、益处及其未来可能的应用或演进，以便更好地理解这项技术及该框架所蕴含的未来，从而塑造这些模型在协同创作情境中的应用，使其切实有用、安全，并最终增强人们及社群的创造能力。

参考文献

[1] Daniel Buschek、Lukas Mecke、Florian Lehmann 和 Hai Dang。2021 年。《设计人机协同创意系统时的九个潜在陷阱》。*arXiv* abs/2104.00358（2021）。[2] Lydia B. Chilton、Anna Kantosalo、Q. Vera Liao、Mary Lou Maher、Charles Martin、Michael Muller 和 Greg Walsh。2023 年。《生成式人工智能与人机交互》 |CHI 2023研讨会。https://generativeaiandhci.github.io[3] IBM Design。“无日期”。《企业设计思维》。https://www.ibm.com/design/approach/design-thinking/[4] Anna Kantosalo 和 Tapio Takala。2020 年。《人机协同创造的五个 C——经典创造力视角的更新》。发表于国际创新计算与云计算会议。[5] A. Baki Kocaballi。2023 年。《对话式人工智能驱动设计：作为设计师、用户和产品的 ChatGPT》。https://doi.org/10.48550/ARXIV.2302.07406[6] Jieun Kwon、Younghyun Choi 和 Yura Hwang。2021 年。《企业设计思维：大型企业以用户为中心的设计流程研究》。*Designs* 5，3（2021）。https://doi.org/10.3390/designs5030043[7] Zhicheng Lin。2023 年。《为何以及如何在学术生活中拥抱 ChatGPT 等人工智能》。https://doi.org/10.31234/osf.io/sdx3j[8] OpenAI。2023 年。《ChatGPT：优化用于对话的语言模型》。https://openai.com/blog/chatgpt/[9] Natalie Staudacher。2023 年。《ChatGPT——发行说明》。https://help.openai.com/en/articles/6825453-chatgpt-release-notes

附录

一个 CHAI-DT“HILLS”提示词

我们正在进行一项“HILLS”设计思维练习。

活动说明：Hill 是以用户为中心的意图陈述，整个团队都可以围绕它凝聚起来，从而朝着同一个方向努力。Hill 描述的是特定用户能够做什么，而不是具体实现方式。它们为团队提供所需的创意空间，帮助团队提出突破性创意，而无需详细需求。请在项目或计划开始时，在确定用户的真实需求之后撰写 Hills。可以把它看作团队使命宣言。采用 Hills 能确保整个团队围绕为用户做正确的事达成一致。它们旨在让所有队友——无论身处不同专业领域、跨越海洋还是处于不同时区——都能步调一致地协作，为用户打造最佳体验。

一个 hill 是团队共同撰写并达成一致的书面陈述。它由三个部分组成，十分简单：

Who（用户是谁）这一部分非常清晰具体，明确指出我们关注的用户类型，例如技术销售人员、业务分析师或人力资源经理。

What（要做什么）是第二部分。我们之前称之为用户赋能。它指的是一项具体任务，或用户能够完成的具体事项。Hill 中的 What 可能是部署某项内容、创建某样东西，或运行某种分析。也就是说，What 是一项具体的赋能，让你在 Who 中提到的用户能够做到某件事。

最后但同样重要的是，Hill 还包括 Wow——Who、What、Wow。这是一项具体市场价值或差异化陈述。它将 What 延伸至令人由衷欣喜的境界，而我们知道，如今与用户合作时，所有工作都必须达到这一点。如果你把自己的 Hill 念给项目发起方用户听（你应该这么做），对方应该会回应：“太棒了，我想要这个。我需要这个。什么时候能给我？”

因此，每个愿景陈述都包含三个部分：谁、什么，以及惊喜之处。

优秀愿景陈述示例：“在选定的产品类别中，请求者可以通过自然语言对话来查找与其搜索查询相匹配的产品。”

本活动的指令：1. 创建“谁”列表 2. 创建“什么”列表 3. 创建“惊喜之处”列表。4. 针对每个部分尽可能多地提出想法，并迅速与队友分享。借鉴他人的想法，但要重数量、轻质量，并避免偏离主题去讨论功能或实现细节。5. 运用你们关于“谁”“什么”和“惊喜之处”的想法，构建愿景陈述。

相关活动情境：大型零售连锁企业 RetailInc 与 IBM 携手改进其库存管理系统。现有系统已经过时，且经常导致缺货和库存积压，造成销售损失和资源浪费。IBM 为合作项目带来数据分析和人工智能方面的专业知识，而 RetailInc 则提供其丰富的销售和库存水平数据。双方旨在共同打造一个新系统，能够准确预测销售趋势，并实时优化库存水平，从而提高零售连锁企业的效率和利润。

鉴于上述情境，请执行练习的第 1 步。

收稿日期：2023年2月20日；录用日期：2023年3月9日