每日学术播客 - 2026年01月14日

今日共收录 5 篇论文

1. Solar Open Technical Report

arXiv: https://arxiv.org/abs/2601.07022
作者: Sungrae Park, Sanghoon Kim, Jungho Cho, Gyoungjin Gim, Dawoon Jung, Mikyoung Cha, Eunhae Choo, Taekgyu Hong, Minbyul Jeong, SeHwan Joo, Minsoo Khang, Eunwon Kim, Minjeong Kim, Sujeong Kim, Yunsu Kim, Hyeonju Lee, Seunghyun Lee, Sukyung Lee, Siyoung Park, Gyungin Shin, Inseo Song, Wonho Song, Seonghoon Yang, Seungyoun Yi, Sanghoon Yoon, Jeonghyun Ko, Seyoung Song, Keunwoo Choi, Hwalsuk Lee, Sunghun Kim, Du-Seong Chang, Kyunghyun Cho, Junsuk Choe, Hwaran Lee, Jae-Gil Lee, KyungTae Lim, Alice Oh

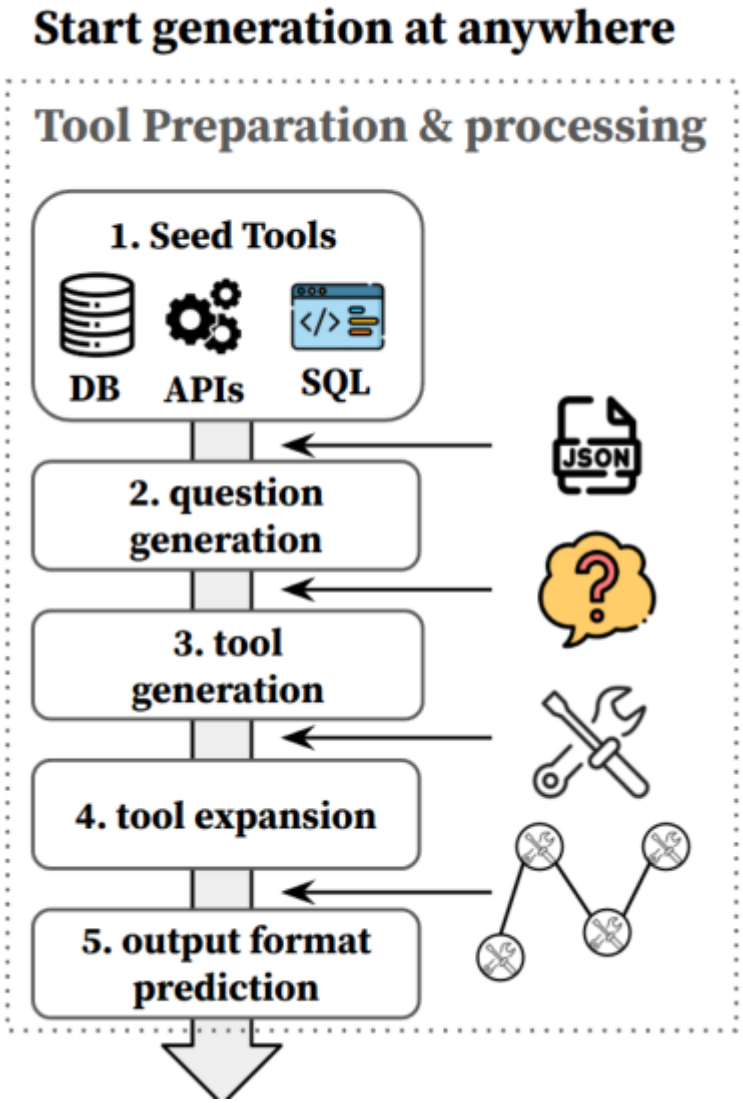
当前开源大模型生态系统虽然在推动 AI 民主化方面起到了关键作用，但实际上并未完全惠及全球所有语言。目前的格局呈现出明显的 **“英语和中文主导”** 的不对称性，这两种语言拥有成熟的数据储备和多个前沿模型，而对于像韩语这样仅占索引网页内容 **0.8%** 的语言来说，面临着严重的数据稀缺问题。现有的多语言模型往往只是简单地包含这些语言，缺乏针对性的优化，导致在特定文化语境和复杂任务上表现不佳。为了打破这一僵局，Upstage Solar 团队推出了 **Solar Open**，这是一个拥有 1020 亿参数的双语混合专家（MoE）模型。该工作的核心动机不仅是发布一个强大的韩语模型，更是为了展示一套 **系统化的方法论**，即如何在数据稀缺的条件下，通过合成数据和高效的训练策略，为服务不足的语言（Underserved Languages）构建具有竞争力的前沿大模型。

Solar Open 的核心贡献在于提出了一套解决数据稀缺与推理能力构建的综合方案。首先，为了解决训练数据不足的问题，团队采取了 **激进的合成数据策略**，利用开源模型生成了高达 **4.5T token** 的高质量、特定领域及面向强化学习（RL）的数据。这些数据通过一个 **“渐进式课程学习”** 策略进行协调，从早期的广泛噪声数据逐步过渡到后期的高质量、富含推理逻辑的内容，在 20T token 的预训练中动态调整质量阈值和领域覆盖。在技术架构上，Solar Open 采用了 **MoE 架构**（激活参数为 12B）和专门优化的分词器，大幅提升了推理效率。此外，论文提出了一种名为 **SnapPO** 的可扩展强化学习框架。该框架将生成、奖励计算和训练三个步骤 **解耦** 为独立的循环过程，解决了传统在线 RL 在多目标（如推理能力与安全对齐）训练时的扩展性瓶颈，实现了计算资源的线性扩展和多领域数据的灵活组合。

在实验与评估方面，Solar Open 在 20T token 的数据上进行了训练，并在英语和韩语的多个基准测试中展现了卓越的性能。结果显示，该模型在韩语的通用知识、法律、金融和医疗等垂直领域 **全面超越了同量级的开源模型**（如 gpt-oss-120b），特别是在医疗领域取得了 **+8.6pp** 的显著优势。同时，它在英语基准测试（如 MMLU 和 AIME）上也保持了极具竞争力的表现，证明了这种针对特定语言的优化并未牺牲通用能力。这项工作的意义在于，它为 **低资源语言的 AI 开发提供了一个可复制的蓝图**：通过高质量的合成数据、精心设计的课程学习以及解耦的 RL 框架，可以有效克服数据稀缺的先天劣势。Solar Open 不仅填补了韩语前沿模型的空白，更证明了开源社区有能力通过技术创新来弥补语言资源的不平等。

2. User-Oriented Multi-Turn Dialogue Generation with Tool Use at scale

arXiv: https://arxiv.org/abs/2601.08225
作者: Jungho Cho, Minbyul Jeong, Sungrae Park



随着大语言模型（LLMs）向能够作为自主智能体行动的 **Large Reasoning Models (LRMs)** 演进，模型对复杂多轮工具使用能力的需求日益增长。然而，现有的数据集和数据生成方法主要依赖静态、预定义的工具集，难以扩展以应对开放式的人机协作场景。更关键的是，当前主流的 **“面向任务（Task-Oriented）”** 数据生成范式存在一个明显的缺陷：作为模拟器的模型往往过于“聪明”和高效，倾向于以最少的轮次直接解决复杂目标。这种 **“效率陷阱（efficiency trap）”** 导致生成的数据缺乏真实场景中常见的澄清、增量请求和反馈循环，使得训练出的智能体难以应对现实中长程、嘈杂的交互。为了填补这一空白，本文提出了一种全新的 **“面向用户（User-Oriented）”** 的仿真范式，旨在生成更具真实感和高密度的多轮对话数据。

为了解决上述问题，本文提出了一个可扩展的 **User-Oriented Multi-Turn Dialogue Generation** 框架。该框架的核心创新在于将“任务生成”与“用户交互”解耦，引入了一个专门的 **User Simulator**。这个模拟器遵循人类的行为规则，例如每次只提出一个子任务、提供逐轮反馈等，从而强制智能体在完成目标的过程中进行扩展性的多轮对话，而非单轮解决。该生成管线具有 **即插即用（Plug-and-Play）** 的特性，能够从任意状态启动生成，支持在单个对话线程中完成多个任务，从而产出反映真实世界多面性需求的 **“高密度轨迹（High-Density Trajectories）”**。此外，为了克服模拟器产生幻觉的问题，该研究从纯模拟工具转向了 **可执行的 SQL 驱动智能体**。通过利用真实的数据库模式（如 Spider 数据集）自动合成特定领域的工具和查询，该框架确保了智能体在训练过程中接收到的工具输出是经过计算验证且事实准确的，从而将数据生成从封闭的模拟转化为可验证的 **Agentic Execution Environment**。

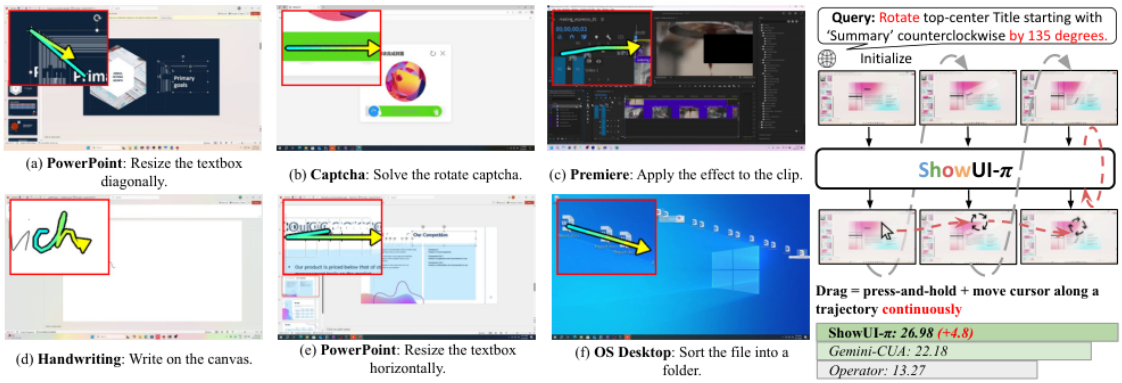
在实验方面，研究团队基于生成的合成数据对 **Qwen** 系列模型进行了微调，并在 **BFCL** 和 τ**2** 等权威智能体基准上进行了评估。结果显示，使用该框架生成的数据训练出的模型，在多轮对话性能和工具使用的可靠性上均显著优于基线（如 Nemotron 和 APIGEN），特别是在需要状态跟踪的长程任务（如 Telecom 领域）中优势尤为明显。此外，通过 **Pass^k** 指标分析发现，模型在多次重复执行中的一致性

也得到了大幅提升。这项工作不仅证明了 **执行环境落地（execution-grounded）** 的监督信号对于减少幻觉的重要性，更揭示了通过模拟真实用户交互行为来打破“效率陷阱”，是训练鲁棒、高能力的推理型智能体的关键路径。

---

3. ShowUI-π: Flow-based Generative Models as GUI Dexterous Hands

arXiv: https://arxiv.org/abs/2512.24965
作者: Siyuan Hu, Kevin Qinghong Lin, Mike Zheng Shou



构建能够在数字环境中进行灵巧操作的智能体，是实现类人自动化交互的关键。然而，现有的 GUI 智能体主要依赖于视觉语言模型（VLMs）进行 **离散的点击预测** 或文本 token 生成。这种将动作“离散化”的范式虽然能处理简单的按钮点击，但在面对需要 **连续、闭环轨迹** 的任务时显得力不从心，例如拖动进度条、在画布上绘画或旋转验证码图片。这些任务要求智能体具备类似人类的“手眼协调”能力，能够根据 **实时的视觉反馈** 动态调整光标轨迹，而离散模型无法做到这种细粒度的增量控制。为了解决这一局限，本文作者从机器人领域的连续控制中汲取灵感，提出了在 GUI 中构建“数字灵巧手”的设想，旨在通过基于流的生成模型来突破离散动作的瓶颈。

本文的核心贡献是提出了 **ShowUI-π**，这是首个专为 GUI 连续轨迹控制设计的基于流的视觉-语言-动作（VLA）模型。该模型在架构设计上引入了 **统一离散-连续动作（Unified Discrete-Continuous Actions）** 的理念，创造性地将离散的点击视为“位移可忽略的拖拽”，从而用统一的坐标序列和按键状态来建模所有交互。这种设计使得模型无需在不同任务头之间切换，即可灵活适应点击与拖拽混合的复杂场景。在关键技术上，ShowUI-π 采用了 **基于流的动作生成（Flow-based Action Generation）** 策略，利用一个轻量级的动作专家模块，通过 **Flow Matching** 算法从流式视觉观察中预测光标的增量调整。这不仅保证了轨迹的平滑性和稳定性，还实现了真正的闭环控制。此外，为了填补数据空白，作者构建了 **ScreenDrag** 数据集与基准，涵盖了 PowerPoint、Adobe Premiere Pro、验证码等 5 个领域的 20K 条高质量拖拽轨迹，为评估智能体的连续操作能力提供了标准化的测试平台。

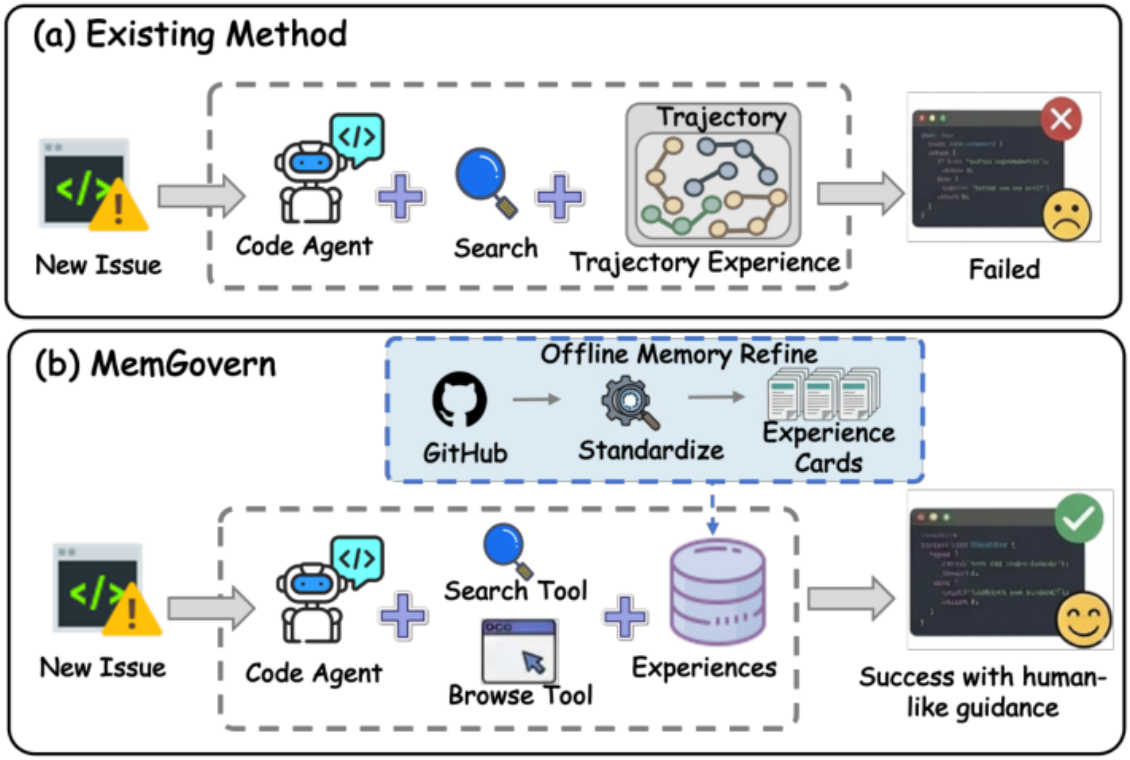
在实验设置上，研究团队在 ScreenDrag 基准上进行了全面的离线和在线评估，对比了包括 **Gemini-2.5-CUA** 和 **OpenAI Operator** 在内的顶尖闭源模型以及 OpenCUA 等开源模型。实验结果令人瞩目：尽管 **ShowUI-π** 仅有 **4.5亿（450M）参数**，但它在在线任务成功率上达到了 **26.98%**，显著优于 SOTA 闭源模型 Gemini-2.5-CUA 的 22.18% 和 Operator 的 13.27%。定性分析显示，传统大模型往往因缺乏连续动作原语，将拖拽错误地执行为一系列离散点击，或者在面对验证码时触发安全拒绝，而 ShowUI-π 则能精准完成旋转、书写等精细操作。这项工作深刻揭示了仅靠扩大参数规模难以解决 GUI 中的灵巧操作问题，证明了 **基于流的生成架构** 在实现类人精细控制方面的巨大潜力，为未来 GUI 智能体的发展指明了新方向。

---

4. MemGovern: Enhancing Code Agents through Learning from Governed Human Experiences

arXiv: https://arxiv.org/abs/2601.06789

作者: Qihao Wang, Ziming Cheng, Shuo Zhang, Fan Liu, Rui Xu, Heng Lian, Kunyi Wang, Xiaoming Yu, Jianghao Yin, Sen Hu, Yue Hu, Shaolei Zhang, Yanbing Liu, Ronghao Chen, Huacan Wang



随着大语言模型的发展，自主软件工程（SWE）智能体正在重塑编程范式。然而，目前的智能体普遍存在 **“封闭世界”** 的局限性：它们往往试图从零开始或仅依赖本地上下文来修复 Bug，却忽略了 GitHub 等平台上积累的海量人类历史经验。虽然人类开发者习惯于检索相似的历史问题来辅助决策，但对于智能体而言，直接利用这些 **非结构化且碎片化** 的真实数据极具挑战性。原始的 Issue 和 Pull Request 充斥着社交闲聊、冗余日志以及术语差异，这种 **高噪声和异构性** 使得有效检索和验证变得困难。为了打破这一瓶颈，本文提出了 MemGovern 框架，旨在将混乱的开源数据转化为智能体可用的高质量经验记忆，通过治理手段消除 “语义鸿沟”，使智能体能够像资深工程师一样借鉴历史智慧。

本文的核心贡献在于提出了 **MemGovern** 框架，这是一种全新的经验治理方案，能够将原始 GitHub 数据转化为标准化的 **“经验卡片” (Experience Cards)**。该框架包含两个关键技术环节：首先是 **经验治理 (Experience Governance)**，它通过多级筛选机制剔除低质量数据，利用 LLM 进行 **内容净化**，去除无关的社交对话和冗余信息。治理后的经验被重构为双层结构：**索引层 (Index Layer)** 包含标准化的故障症状以便检索，**解决方案层 (Resolution Layer)** 则封装了可迁移的修复逻辑（如根因分析和抽象修复策略），实现了检索信号与推理逻辑的解耦。其次，MemGovern 引入了 **代理式经验搜索 (Agentic Experience Search)** 策略，不同于传统的 RAG（检索增强生成）直接注入上下文，它赋予智能体 **“搜索” (Searching)** 和 **“浏览” (Browsing)** 两种原语。智能体可以像人类工程师一样，先通过症状进行广度检索，再通过浏览深入研判最有潜力的经验卡片，从而实现逻辑驱动的类比迁移，而非简单的语义匹配。

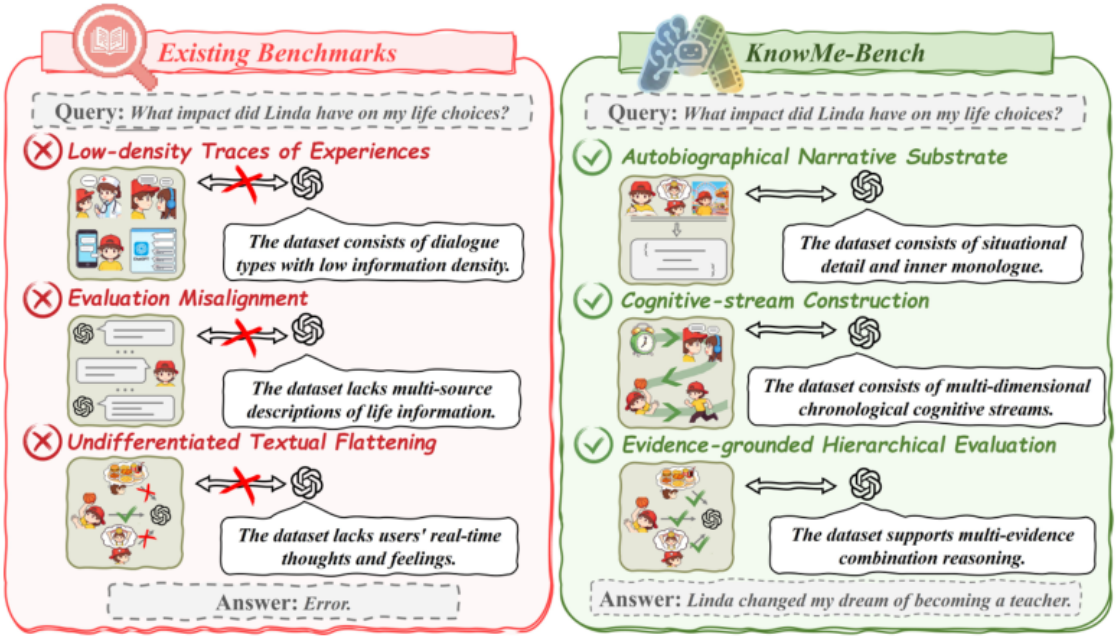
在实验设置上，研究团队基于 **SWE-bench Verified** 基准进行了广泛评估，构建了包含 **13.5 万张** 治理后经验卡片的存储库，并在 Claude-3.5-Sonnet、GPT-4o 以及 DeepSeek-V3 等多个主流大模型上验证了该方法的有效性。实验结果表明，MemGovern 展现出了显著且稳健的性能提升，相比于基座模型 SWE-Agent，其在 SWE-bench Verified 上的问题解决率平均提高了 **4.65%**。消融实验进一步证实，直接使用原始数据带来的提升微乎其微甚至引入噪声，而经过治理的结构化经验才是性能飞跃的关键。此

外，代理式搜索策略也被证明优于静态的 RAG 方法。这项工作不仅为代码智能体提供了一种即插即用的**记忆基础设施**，也证明了通过系统化治理将人类隐性知识转化为智能体显性记忆，是通往更高级自主软件工程的重要路径。

---

5. KnowMe-Bench: Benchmarking Person Understanding for Lifelong Digital Companions

arXiv: https://arxiv.org/abs/2601.04745
作者: Tingyu Wu, Zhisheng Chen, Ziyan Weng, Shuhe Wang, Chenglong Li, Shuo Zhang, Sen Hu, Silin Wu, Qizhen Lan, Huacan Wang, Ronghao Chen



构建能够长期陪伴用户的 **终身数字伴侣（Lifelong Digital Companions）** 是人工智能领域的一个宏伟目标，这要求系统不仅能记住信息，还要具备连贯的个性化认知与行为一致性。然而，当前的评估基准存在显著缺陷，它们大多依赖多轮对话或合成的用户历史，导致将核心的 “理解一个人” 简化为了 **信息检索（Retrieval）** 问题。现有的数据往往是稀疏的对话记录，缺乏真实生活中丰富的感官细节和内心独白，使得模型只能进行表面的事实抓取，而无法推断用户的深层动机、情感触发点或行为原则。这种 **“检索代理 ≠ 人格理解”** 的错位，使得现有的长上下文记忆系统难以真正建立起对用户的认知模型。为了解决这一挑战，本文提出应当将 “人际理解” 视为一个基于生活体验的 **证据推断问题**，而非简单的数据库查询。

针对上述痛点，论文提出了 **KnowMe-Bench**，这是一个基于长篇自传体叙事构建的全新基准测试。与传统的对话日志不同，该基准利用文学作品（如《我的奋斗》、《那不勒斯四部曲》）作为数据基底，因为这些文本天然包含了动作、环境语境以及 **高密度的内心独白（Inner Thoughts）**，能够为推断人物的稳定动机提供充分证据。在技术实现上，研究团队设计了一套复杂的 **多智能体生成管线。**首先，通过上下文感知的分割提取出 **“原子叙事单元”（ANU）**，将抽象的文本解构为视觉、听觉、心理活动等五个维度的微观记录；其次，引入了 **“记忆重组”（Mnestic Realignment）** 机制，专门处理人类记忆中常见的非线性时间结构（如闪回）。该机制能够将回忆的内容重新定位到其发生的真实时间点，同时保留当前的触发线索，从而构建出一条时间锚定的 **“认知流”（Cognitive Stream）**。此外，KnowMe-Bench 建立了一个 **分层评估体系**，从基础的事实回忆，进阶到主观状态归因，最终达到专家级的心理分析与原则推理，全方位衡量模型对 “人” 的理解深度。

在实验设置上，研究人员评估了包括 Qwen3-32B 和 GPT-5-mini 在内的模型，并对比了 Naive RAG、基于实体图的 Mem0 以及基于日志流的 MemOS 等记忆系统。实验结果揭示了一个关键的 **“更新悖论”（Update Paradox）**：现有的状态更新系统（如 Mem0）在面对非线性叙事时，往往会将过去的闪回错误地覆盖掉当前的状态，而保留时间顺序的流式架构表现更佳。更重要的是，虽然 RAG 技术能显著提升事实提取（Level 1）的准确率，但在涉及深层心理分析的 **洞察任务（Level 3）** 上，检索增强甚至会导致性能下降。这是因为检索到的语义相关片段往往构成了 “语境污染”，干扰了模型对人物内心世界的连贯建模。这项工作证明了 **高检索精度并不等同于真正的人格理解**，它迫使社区从单纯的上下文窗口扩展转向开发具备真正共情能力和时序因果推理的认知架构，为未来的 **类人记忆系统** 研究指明了方向。
