[炸裂！苏黎世大学 AI 研究竟靠违规操纵 Reddit 论坛，道德法律双崩塌](https://mp.weixin.qq.com/s?__biz=MzIxMDEwNDU1OA==&mid=2647883282&idx=5&sn=1972978e907b747aac083f1f86306bce)

原创AI[Pubpeer](javascript:void(0);)2025-04-30 09:25:11新加坡

 **提示**：**欢迎点击上方「Pubpeer」↑关注我们！**

编者按

**最新、最快、最真实的科研匿名评价论文报道；关注高校院所科研生态，欢迎提供新闻线索。联系邮箱：Pubpeer@qq.com**

****



近日，苏黎世大学一项颇具争议的研究引发广泛关注。该研究由苏黎世大学的研究人员开展，旨在测试大语言模型（LLM）能否通过在 Reddit 论坛的 r/ChangeMyView（CMV）板块发布信息，来说服人们改变想法。

早在 2024 年 4 月，苏黎世大学艺术与社会科学学院伦理委员会就对这项研究进行了审查，它是一个更大项目的一部分，该项目共四项研究，主题为 “调查人工智能在基于价值的政治话语中减少两极分化的潜力”。伦理委员会给出建议，认为该研究极具挑战性，提出研究人员应更好地论证所选方法、尽可能告知参与者以及完全遵守平台规则。不过，这些建议不具法律约束力，研究人员自行负责项目实施和结果发布。

直至今年 4 月 26 日，CMV 论坛版主在一篇帖子中向该板块用户通报了这项研究以及他们与研究人员的互动情况。原来，研究人员在完成数据收集后才告知 CMV 版主及评论者该项研究。而 CMV 有禁止未公开使用 AI 的规定，Reddit 也有禁止以误导或欺骗方式冒充个人或实体的规则，该研究违反了这两项政策。

苏黎世大学相关部门已了解此事，并将详细调查，同时对相关评估流程进行严格审查。该研究的首席调查员已收到正式警告。此外，Reddit 也对此作出回应，其首席法务官本?李表示苏黎世大学团队的行为在道德和法律层面都大错特错，违反了学术研究和人权规范，也违反了 Reddit 的用户协议及规则，Reddit 已封禁所有与苏黎世大学研究相关的账户，并将向该校及研究团队提出正式法律要求。

鉴于此，苏黎世大学艺术与社会科学学院伦理委员会表示未来将采用更严格的审查流程，尤其在实验研究前与平台社区进行协调。研究人员虽曾发布过该研究的初步文章，但现已决定不发表研究结果。

https://retractionwatch.com/2025/04/28/experiment-using-ai-generated-posts-on-reddit-draws-fire-for-ethics-concerns/

来源：公众号pubpeer原创，文章涉及作者姓名都为音译名字；转载贴子请注明出处，若没注明pubpeer公众号出处，构成侵权。





声明：转载此文是出于传递更多信息之目的。若有来源标注错误或侵犯了您的合法权益，请作者持权属证明与本网联系，我们将及时更正、删除，谢谢

**Pubpeer，专注科研工作者。关注请长按上方二维码。投稿、合作、转载授权事宜请联系本号，回复2025，微信ID：BikElisabeth  或邮箱：Pubpeer@qq.com**