

生成式人工智能降低了计算社会科学的研究门槛吗？*

张勇军 著** 张文菲 译***

提要:生成式人工智能正深刻变革着计算社会科学领域,为数据收集与分析的多模态处理开辟了新的可能。这一技术突破对缺乏深厚编程背景的研究者而言影响重大。其一,生成式人工智能可通过自动化代码生成、注释和调试,大幅提高社会科学研究者的工作效率;其二,借助创新的提示工程,研究人员能够深入开展复杂的数据分析;其三,计算社会科学的教學领域也能受益于生成式人工智能工具,尤其是在代码注释和复杂代码解释方面,从而简化了学习过程,使计算社会科学技术更易理解和接触。

关键词:生成式人工智能 计算社会科学 多模态计算 提示词工程

以 OpenAI 的 GPT-4 系列和 Deepseek R1 为代表的生成式人工智能的迅猛发展,正显著重塑计算社会科学(computational social science, CSS)的格局。计算社会科学作为一个跨学科领域,旨在借助自然语言处理(NLP)、计算机视觉等一系列计算工具(Edelmann et al., 2020; Lazer et al., 2020),通过剖析社交媒体、行政记录及其他数字化来源的大规模数据集,推动社会理论研究的

* 本文原文为英文,中文翻译已获得作者本人正式授权。作者感谢纽约州立石溪大学高级计算科学研究所对高性能计算系统、OpenAI API 和 ChatGPT 的支持。

** 张勇军(yongjun.zhang@stonybrook.edu),亚利桑那大学社会学博士,现任美国纽约州立石溪大学(SUNY Stony Brook University)社会学系暨高级计算科学研究所助理教授。研究重点关注政治社会学、组织行为、大数据与计算社会科学等领域。

*** 张文菲,云南大学民族学与社会学学院。

进步。该领域的研究者常面临一项颇具挑战性的任务：如何从文本、图像、音频和视频等多模态数据来源中收集并提炼有价值的见解？例如，有学者(Fan & Zhang, 2022)运用结构主题建模方法，对涵盖65种语言的大规模在线新闻数据集进行分析，探究了全球媒体在新冠肺炎大流行期间对中国的报道情况。在另一项研究中，张(音)等人(Zhang et al., 2023)通过微调自有大型语言模型，分析了数百万条推文，以评估公众对中国的情感倾向。

在生成式人工智能时代到来之前，进行此类分析要求研究者具备高级编程技能，如编写网页抓取脚本、微调预训练的大型语言和视觉模型，以及完成情感分析和图像分类等下游任务。如今，以OpenAI的GPT系列和Meta的LLaMa等为代表的生成式人工智能工具，能让研究者轻松地将多样化的输入转化为连贯的文本输出。更具创新性的是，这些工具具备高度精细的提示词工程(prompt engineering)能力，不仅能简化代码的生成、注释和调试过程，还能使复杂的计算任务变得更加轻松。最新技术进展表明，这些工具可通过创新提示词工程，直接用于数据的收集与分析。

本文首先剖析长期制约计算社会科学研究者获取必要技术技能的主要障碍，接着探讨生成式人工智能工具如何通过降低编程复杂度、简化数据收集与分析流程，助力研究者克服这些障碍。最后，本文还将探讨在计算社会科学领域应用生成式人工智能所引发的认识论风险，并提出相应的应对策略。

一、计算社会科学面临的主要障碍

在过去20年里，计算社会科学作为一个蓬勃发展的领域，在中国和西方社会均取得了显著进展(Lazer et al., 2020; Wu et al., 2024)。在其发展初期，计算社会科学主要依托资源丰富的精英机构。尽管“计算社会科学”这一术语的首次提出者尚无定论，但该领域在2009年因大卫·拉泽等人(Lazer et al., 2009)发表于《科学》(*Science*)期刊的开创性文章而得到广泛认可。在此之前，2007年一批满怀壮志的科学家齐聚一堂，共同探索大数据分析在社会

科学中的应用潜力。会上,哈佛大学政治学学者加里·金(G. King)指出,数字信息的爆炸式增长“将使我们能够更深入地了解社会,最终真正解决影响人类群体福祉的重大问题”(Ledford, 2020)。

20世纪80年代末90年代初,互联网的诞生彻底改变了人类的生活与互动方式。社会生活日益网络化,人与人之间的联系愈发紧密。拉泽等人在2009年《科学》期刊的文章中生动阐述了这一数字化转型的影响:

我们会定期查看电子邮件,无论身处何地都能通过手机通话,刷卡乘坐公共交通工具,使用信用卡购物。我们在公共场所的活动可能会被监控摄像头记录,医疗记录也已实现数字化。我们可能会发布可供任何人访问的博客,或通过社交网络维系友谊。每一笔交易都会留下数字痕迹,这些痕迹汇聚成个人和群体行为的完整图景,具有变革我们对生活、组织和社会认知的巨大潜力。

这一描述表明,我们已身处数字时代数十年。大数据的爆炸式增长为研究者提供了强大的新工具,使其能够借助人们留下的数字痕迹深入研究人类和社会行为。这些数据涵盖大规模档案数据集、行政记录、图像与视频数据等。尽管机遇众多,但计算社会科学作为一个正式的学术领域,其发展仍较为缓慢。早期的创新主要集中于谷歌、脸书、亚马逊、微软等大型科技公司以及政府机构,直到近年来,社会科学的主流学科才逐渐接纳并发表更多基于计算方法的研究。过去十年间,计算社会科学在美国逐渐站稳脚跟。这一发展得益于大规模数据的更广泛获取、计算资源的普及、分析方法的进步以及研究社区的日益成熟。

然而,计算社会科学的发展仍面临数据收集与数据分析这两大主要障碍。

首先,信息与通信技术的迅猛发展在带来海量数据的同时,也造成了技术鸿沟。许多传统社会科学学者缺乏通过网页抓取(web scraping)或应用程序接口(API)访问大规模结构化和非结构化数据的能力。例如,在美国社会运动研究中,社会运动组织如何参与政策制定过程是关键的问题(Maher et al., 2024; Seguin et al., 2023)。一种分析方法是研究国会证词,这要求研究

者运用网页抓取技术,从亚利桑那大学的 ProQuest 国会听证数据库中提取 1946—2016 年所有听证会的元数据,包括委员会信息、听证日期以及近 88 万个证词的详细内容。这些证词涉及超过 91 000 场听证会,包含证人姓名、所属机构和职称等信息。

在生成式人工智能时代之前,从动态网站抓取数据需要拥有较高的技术水平,如编写 Selenium 自动化脚本,运用 Python 或 R 语言进行网页抓取,这对于没有计算机科学背景的学者来说是巨大障碍。即便 API 访问被视作一种更易获取公共网站或社交媒体平台数据的方式,但仍需较强的编程技能。例如,为跟踪美国总统选举期间油管上的相关活动,笔者所在的研究团队开发了一个数据基础模块,借助 YouTube Data API v3,每天从该平台查询并收集数百万个视频的元数据,涵盖观看次数、点赞数、点踩数、评论、缩略图及频道信息。项目实施需要复杂的 Python 脚本来查询 API、筛选内容、管理远程服务器的数据存储并处理突发错误,这些任务对未接受技术训练的社会科学研究者而言极具挑战性。

其次,尽管多模态数据的快速增长为社会科学研究者提供了丰富的研究契机,但有效利用这些资源往往需要借助自然语言处理和计算机视觉等高级计算方法。例如,城市研究者越来越多地运用卫星图像评估贫困和财富状况,尤其是在传统经济数据匮乏的地区。在许多发展中国家,由于缺乏全国代表性的数据,国际组织(如联合国)难以精准制定扶贫政策。为弥补这一缺陷,研究人员利用白天的卫星图像估算经济数据,并取得了良好成效(Burke et al., 2021; Jean et al., 2016)。

传统上,分析大规模图像数据需要具备在 PyTorch 或 TensorFlow 等框架中编写代码的专业能力,以便训练卷积神经网络(CNN)。随着深度学习的发展,尤其是 GoogLeNet、VGG、ResNet 和 Inception 等架构的出现,研究人员能够从图像中提取有价值的特征。近年来,基于变换器(Transformer)的模型(如 Swin Transformer)在计算机视觉任务中表现出色(Liu et al., 2021)。然而,由于标注数据稀缺,社会科学研究者很少从头开始训练大型神经网络。迁移学习提供了一种有效解决方案:预训练模型可用于从新图像中提取低级特征(Wankmüller, 2022)。吉恩等人(Jean et al., 2016)采用这一策略,利用

来自五个非洲国家的数据预测当地经济状况。他们先使用在 ImageNet 上预训练的卷积神经网络,微调该模型以预测卫星图像中的夜间光强度。这种方法借助夜间光强度作为全球经济活动的代理变量,尽管该指标并不完美,但已被广泛应用。在此过程中,模型学会将高维度的输入图像转化为紧凑的特征向量,从而有效预测夜间光强度。随后,这些特征向量与调查数据结合,用于训练岭回归模型,估算经济指标,如经济支出或资产。尽管迁移学习使学者无需从头开始训练深度学习模型,但微调这些模型仍需编写 PyTorch 或 TensorFlow 代码,这在生成式人工智能出现之前构成了显著的技术障碍。

综上所述,社会科学研究者进入计算研究领域的主要阻碍在于,要想完成大规模在线数据收集以及从多模态数据(如文本和图像)中提取关键信息等任务,首先需要编写复杂代码。在生成式人工智能兴起之前,研究人员通常借助 Python Selenium 等工具抓取动态网站,并通过微调预训练模型(如用于文本的 BERT、用于图像的 ResNet 或 Swin Transformer)执行分类任务。这些过程要求研究者深入理解 PyTorch 或 TensorFlow 等框架,给缺乏计算机编程背景的学者带来了较大的学习压力。接下来,本文将阐述如何通过生成式人工智能简化编程复杂性、优化数据收集和分析流程,从而有效降低这些技术门槛。

二、生成式人工智能驱动下的编程

在计算社会科学领域,编程已不再是研究者面临的主要挑战。在生成式人工智能的推动下,以 OpenAI 的 GPT 模型为代表的智能工具,正在重塑研究者与编程之间的关系。编程不应再成为有志于从事计算社会科学研究的学生必须跨越的主要障碍。如今,生成式人工智能工具能够自动生成、注释和调试代码,协助完成各类数据收集与分析任务。这并不意味着研究者无需学习计算社会科学方法,而是表明技术的入门门槛正在降低,使更多研究人员能够投身计算社会科学领域,而不必受编程本身的束缚。

(一)代码生成

生成式人工智能的一项重要进展是其具备生成可用代码的能力。诸如 GitHub Copilot、Code LLaMa 和 ChatGPT 代码解释器等工具,已成为研究者和工程师日常工作中不可或缺的助手,极大地提高了代码生成的便捷性,对编程经验有限的社会科学研究者帮助尤为显著。借助 OpenAI 的 ChatGPT(GPT-4),研究者能够快速生成并优化脚本,完成诸如图像分类任务这类复杂的计算分析。

以笔者的研究项目为例,该项目涉及一个包含超过 200 万张可能描绘社会抗议的图像数据集。这些图像存储在网络附加存储(NAS)中,按国家分目录存放,再进一步按国家代码细分为子目录。此外,项目还引入了 UCLA 抗议图像数据集,其中有 40 764 张图像,11 659 张与抗议相关,每张图像均标注了不同的视觉属性和情感属性(Won et al., 2017)。需要注意的是,尽管训练数据集中的图像已分为训练集和测试集文件夹,但相应标签存放在独立的文件中。这些文件包含多个变量,如文件名和抗议标签,其中文件名与训练集和测试集文件夹中的图像文件名相匹配。项目有两个核心目标:其一,编写 Python 代码加载 UCLA 抗议图像数据集,并微调预训练的 Swin Transformer v2 模型以识别抗议图像;其二,将微调后的模型应用于 200 万张图像数据集的分类。

以下提示词展示了如何利用 ChatGPT 完成上述任务。经对其生成的代码(未在正文中展示)进行深入分析,发现 ChatGPT 能够以清晰且结构化的方式“理解”任务需求。尽管自动生成的代码在细节方面仍有提升空间,但总体而言已具备较高的可用性。

提示词

编写 PyTorch 代码来微调 Swin Transformer 模型 v2。训练图像存于训练集和测试集文件夹中,标签则存于两个独立文件中:train_annot.txt 和 test_annot.txt。需注意,这些文本文件包含 fname 和 protest 等变量。fname 是对应训练集和测试集文件夹中图像的文件名。编写 Python 代码以加载图像和标签数据,并进行 Swin Transformer 模型的微调、评估及保存。最后,用保存的模型对 images_by_country 文件夹中的图像进行推理,

该文件夹的子文件夹以国家名称命名。确保编写 Python 代码来检查这些图像文件的有效性。

ChatGPT 回答

以下是使用 PyTorch 微调 Swin Transformer 模型的 Python 代码。此代码假设您已安装必要的环境和预训练的 Swin Transformer 模型。步骤如下：

1. 读取注释文件并准备数据集。
2. 定义一个自定义数据集类来加载图像和标签。
3. 定义模型、损失函数和优化器。
4. 训练模型。
5. 在测试数据集上评估模型。
6. 保存微调后的模型。
7. 使用保存的模型对来自 images_by_country 的新图像进行推理，并检查图像文件的有效性。

首先，确保已安装所需的库：

```
pip install torch torchvision timm pandas pillow
```

以下是完整的 Python 代码……^①

(二) 代码注释

生成式人工智能工具能够借助代码注释，切实有效地辅助社会科学研究者。初始生成的代码输出中，常常存在复杂或难以理解的代码片段。对于编程新手而言，这样的代码结构与逻辑极具挑战性。研究者面对这些代码时，可能无法迅速把握函数逻辑及参数设置的具体含义。

为弥补这一不足，研究者可借助 ChatGPT 等代码解释器，以逐行注释的方式梳理代码结构。例如，用户可向 ChatGPT 提出如下请求：

我从初步输出中获取了这段代码，然而作为编程新手，我很难理解

^① 限于篇幅，未予展示，读者可在各类人工智能大模型上自行操作或来信索取。——编辑部

其结构和逻辑。能否逐行解释这段代码,并为每个部分添加注释,说明不同参数的作用?这样我就能更好地理解每一部分代码的功能,并依据自身具体需求进行调整。

通过这种方式,ChatGPT 不仅能够分步骤阐释代码的工作原理,还能帮助用户理解编程中的关键概念,从而提升代码阅读与修改能力。这种逐行注释的解释,不仅有助于即时理解代码功能,更重要的是,它能在潜移默化中传授编程知识,使研究者逐渐具备独立修改和编写代码的能力。

具体提示词示例如下:

提示词

我不理解上面生成代码中的 `train` 函数。请逐行注释该代码,并详细解释给我。确保解释这些函数中的参数。

ChatGPT 的回答

当然!以下是 `train` 函数的详细注释,逐行分解并解释其功能……^①

(三)代码调试

当从 ChatGPT 获取的初始代码输出无法完全满足需求时,采用更具针对性和具体性的提示进行迭代是明智之举。这一反复修改的过程对优化代码、使其更好地适应特定目标与约束条件至关重要。通过逐步调整与细化,可使生成的代码更高效,更符合研究预期的理想效果。

例如,最初生成的代码可能无法加载在自然语言处理任务中广泛应用的 HuggingFace 的 `transformers` 库。此外,初始代码可能包含需手动实现的训练循环,而这一过程可借助 `transformers` 库中的内置 `Trainer` 类函数来大幅简化。识别这些优化机会需要一定的专业知识,不过一旦明确,ChatGPT 就能根据提示轻松进行相应修改。

以下是用于引导进一步代码开发的提示词示例:

^① 限于篇幅,未予展示,读者可在各类人工智能大模型上自行操作或来信索取。——编辑部

提示词

使用 HuggingFace 的 Transformers 库修改上述代码,并确保使用 trainer 函数,而不是 PyTorch 代码。

ChatGPT 的回答

这是修改后的脚本版本……^①

笔者使用一台配备了两块 GPU 以及 128 核心 CPU 的工作站运行上述代码,返回如下错误:

代码错误: ImportError: cannot import name “SwinProcessor” from “transformers”

笔者使用以下提示词向 ChatGPT 请求修复代码:

提示词

运行你的代码后,我遇到以下错误: After running your code, I got the following error: ImportError: cannot import name “SwinProcessor” from “transformers”. 请修复它。

ChatGPT 回复

对于之前造成的混淆,我深表歉意。根据我最新的更新,transformers 库中并不包含 SwinProcessor。相反,您应使用为图像模型设计的特征提取器。接下来,我们通过使用 SwinForImageClassification 并搭配合适的特征提取器来修正代码……(下略)

经过多轮迭代调试,微调后的 Swin Transformer 模型在识别图像是否与抗议相关的任务中准确率达到了 94%。训练和推理过程的优化代码已在附录^②中详细列出。

^① 限于篇幅,未予展示,读者可在各类人工智能大模型上自行操作或来信索取。——编辑部

^② 限于篇幅,未予展示,读者若有需求可来信索取。——编辑部

三、生成式人工智能驱动下的多模态数据处理

生成式人工智能不仅能够推动代码生成,还可借助有效的提示词工程,直接开展多模态数据的收集与分析工作。提示词作为输入到大型语言模型的文本指令,其作用是引导模型执行特定任务,使研究者无需进行传统的模型微调,就能直接完成诸多下游应用。而提示词工程是指对这些提示词进行精心设计与优化,以确保模型输出准确且相关的内容。

(一)数据检索

在数据检索方面,生成式人工智能不仅能够生成数据抓取脚本,还能直接进行网页搜索。例如,OpenAI 的部分版本(如 GPT-4o-search-preview)具备浏览功能,可让研究者在生成响应前获取互联网上的最新信息。这一功能对于获取最新数据、监测实时趋势或验证当前事实至关重要,尤其在快速发展的研究领域。当然,学者在使用这些工具时,需警惕生成式人工智能可能出现的“幻觉”现象,即输出内容可能存在不准确或不真实的问题。

以下是一个运用 GPT-4o 搜索功能检索近期美国抗议活动数据的实例:研究者作为社会运动领域的专家,通过网页搜索识别出前一天美国发生的所有抗议事件,记录每个事件的来源网址、内容概要,并将关键信息整理成结构化表格。此方法凸显了生成式人工智能在动态数据收集和 Information 组织方面的实用性。

提示词

你是社会运动领域的专家学者,正通过网页搜索来识别昨天在美国发生的所有抗议事件。请通过保存网址、内容,并总结这些网站中有关事件的关键信息来检索这些事件,输出需整理成结构化表格。

ChatGPT 的回答

以下是关于 2025 年 × 月 × × 日在美国发生的显著抗议事件的结构

化摘要……^①

学者在使用生成式人工智能的网页搜索和文本摘要功能时，结果可能并不全面。然而，这些工具为研究者，尤其是不擅长编写复杂 Python 代码的研究者，提供了宝贵的起点。生成式人工智能能够快速提取和整理信息，使没有编程背景的研究者也能高效获取所需数据。

(二) 数据解读

随着生成式人工智能技术的迅猛发展，这些模型处理多模态数据（如文本、图像、音频和视频）的能力日益精湛。以 GPT-4 模型为例，它拥有视觉识别能力。比如，有一张来自公共广播电台的图片，展示的是 2019 年华盛顿的女性大游行，成千上万的参与者在游行中呼吁女性权利。把这张图片上传至 GPT-4 并给出针对性提示后，该模型就能高效提取关键信息，如人群规模、口号标语、组织标识以及情绪氛围。

这一示例不仅展现了生成式人工智能解读复杂视觉数据的能力，还凸显了其作为社会科学研究工具的巨大潜力。借助这些工具，研究者无需具备技术专长，便能从多模态数据中生成有洞察力的分析。

提示词

请用简洁的句子描述这张图片，告诉我这是否与抗议有关，并展示你的推理过程。

ChatGPT 的回答

这张图片展示了一群参与游行的人，他们举着写有“END WOMEN'S WAGE DISC.”和“2019 WOMEN'S MARCH ON WASHINGTON”的大幅横幅，表明这是 2019 年华盛顿女性大游行。参与者手持标语牌，部分人戴着象征女性大游行运动的粉红色帽子，前景有人使用扩音器，这暗示着演讲或口号是活动的一部分。显然，这与抗议相关，因为女性大游行是

^① 限于篇幅，未予展示，读者可在各类人工智能大模型上自行操作或来信索取。——编辑部

倡导女性权利和性别平等的著名活动,标语牌、大幅横幅及集体游行行动都与抗议活动高度契合。

提示词

总结这张图片,并告诉我其情感、内容和人群规模。

ChatGPT 的回答

这张图片抓拍了 2019 年华盛顿女性大游行的瞬间,参与者举着反对女性薪酬歧视的横幅。画面在情感上展现出团结和积极行动的氛围,旨在推动性别平等和社会正义。标语和符号,如粉红色帽子和举起的拳头,传达了变革诉求以及游行参与者之间的团结。从内容上看,照片展示了不同年龄和民族背景的参与者,一些显眼的标语传递着赋权和抗议信息。横幅周围可见数十人,尽管照片未展示游行全貌,但参与者总数可能达数千人,这种情况在这类活动中较为常见。

(三)数据分析

对于需要处理大量图像的研究者而言,将 OpenAI 的 GPT-4 及其视觉 API 相结合,是强大的研究工具。虽然使用该服务需要订阅,但这一投入能让研究者解锁 GPT-4 在数据分析方面的先进能力。研究者借助这些功能,可高效完成基础数据分析任务,显著提高研究效率。

以笔者对“集体行动动态数据集”的研究为例,GPT-4 不仅能够分析该数据集,还能将结果可视化。把数据集的一部分上传至 GPT-4 并给出适当提示词,模型就能够提取复杂的数据可视化图表,生成一张详细地图,展示美国各地抗议活动的地理分布差异。这一可视化成果不仅揭示了抗议活动的地域分布规律,还为社会运动研究提供了直观参考。值得注意的是,用户可以选择限制 GPT-4 在模型训练中使用其数据,以确保数据控制和保密性。

提示词

请描述 CSV 文件并总结数据模式。

ChatGPT 的回答

我首先需要加载并检查 CSV 文件的内容,以便为您提供文件描述并总结数据模式。请允许我为您进行操作……^①

提示词

请提供各州事件的发生频率。

ChatGPT 的回答

各州事件的发生频率如下……^②

四、使用生成式人工智能的认识论风险

毫无疑问,生成式人工智能通过降低编码复杂性,大幅降低了计算社会科学的入门门槛。然而,其在社会科学研究中的应用也带来了重要的认识论风险,尤其是在缺乏批判性反思而盲目采用大型语言模型的情况下。正如贝尔(Bail, 2024)所警示的,盲目使用生成式人工智能可能会产生“科学垃圾”。尽管一些研究表明,像 ChatGPT 这样的工具在特定任务(如文本摘要)中的表现优于部分人类工作者(Gilardi et al., 2023),但系统性评估也揭示了这些工具存在明显局限。例如,郝(音)等人(Hao & Zhang, 2025)指出,大语言模型生成的输出对提示词工程极为敏感,这种敏感性会引发数据注释过程中的可重复性和有效性问题。

为应对这些挑战,计算机科学家提出了多种提示词策略,如上下文学习、思维链、思维树和角色扮演提示(Wang et al., 2023; Yao et al., 2023),以提升模型性能。然而,对于这些技术如何在计算社会科学中得到有效整合,目前仍缺乏广泛共识。虽然生成式人工智能工具能够在数据收集和分析中协助学者实现编码工作的自动化,但由于模型幻觉等现象,其可靠性与可复制性仍存在不确定性。值得注意的是,即便在相同的提示和参数设置下,大语言模型的输出也可能有所不同,这进一步加剧了社会科学领域业已存在的“复

① 限于篇幅,未予展示,读者可在各类人工智能大模型上自行操作或来信索取。——编辑部

② 限于篇幅,未予展示,读者可在各类人工智能大模型上自行操作或来信索取。——编辑部

刻”危机。

此外,越来越多的研究者依赖 OpenAI、Google 和 Anthropic 等头部人工智能公司开发的专有大语言模型,这也带来了额外挑战。这些模型可能随时间出现“模型漂移”,且架构通常不透明,进一步阻碍了结果的“复刻”(Ollion et al., 2024)。同时,使用敏感的社会科学数据来训练这些模型,还会引发数据隐私和伦理争议等问题。这些问题凸显了在社会科学研究中使用生成式人工智能时,必须进行严格的批判性审视和方法论审查。

在此背景下,虽然生成式人工智能降低了计算社会科学的入门门槛,但研究者在使用这些工具时必须秉持批判性思维并保持足够的谨慎。对这些工具背后计算方法的基本理解,涵盖其优点与局限性,依旧不可或缺。缺乏这种理解会提升误用风险,进而引发经典的“垃圾进,垃圾出”问题,也就是说,不当的输入或工具的错误使用可能产生误导性或无效的结果。

基于郝(音)等人(Hao & Zhang, 2025)的研究,笔者提出以下建议,以帮助研究者更高效地运用生成式人工智能进行直接的数据收集与分析:

其一,运用生成式人工智能的网页搜索功能检索数据时,务必提示模型保存 URL 及来源信息。这既有助于研究人员验证所获内容的可靠性,又能确保数据收集过程的透明性。

其二,直接借助生成式人工智能工具进行数据注释时,研究者应构建一个开发数据集,用以系统地测试不同的提示策略。研究人员要考虑不同的参数设置与提示词方式,通过该开发数据集找出最佳的提示词方法。此方法可助力研究者在多次试验中逐步优化提示效果,确保模型输出结果的稳定性与准确性。

其三,面对复杂任务,研究者应尝试将其分解为更简单的子任务,以此提高模型性能。这是因为,生成式人工智能工具通常在处理简单任务时表现更佳。

其四,研究者可明确要求生成式人工智能模型输出中间推理步骤,这有助于提升模型推理过程的透明度,还能让研究者更易于监控和理解模型的逻辑流。

其五,研究者应优先选用开源模型,如 Meta Llama 系列和 DeepSeek R1 模

型,以保护数据隐私并增强结果的可复制性。对于担心敏感数据泄露的研究人员而言,该策略尤为关键。

其六,交叉验证是任何计算社会科学研究的重要环节。研究者不应盲目信赖生成式人工智能的输出,而应始终构建包含人工标注的评估数据集,用于评估模型性能,并将关键指标公开报告给学术同行。

总之,这些建议凸显了在社会科学研究中运用生成式人工智能时,保持方法论严谨性和伦理意识的重要意义,尤其是在生成式人工智能逐渐融入研究流程的背景下。只有在清晰理解其优势与局限的基础上,研究者才能充分利用生成式人工智能,同时避免可能出现的误用和偏差。

五、结 论

本文旨在阐明生成式人工智能的出现如何推动计算社会科学方法的普及,使编程能力较弱的社会科学研究者也能够借助先进的计算工具开展研究。

这一技术飞跃的关键在于,编程以往是数字人文学科和计算社会科学研究的主要入门障碍,如今却被重新定义为促进研究的工具,而非必须跨越的门槛。

正如贝尔(Bail, 2024)、戴维森(Davidson, 2024)所指出的,生成式人工智能已成为社会科学研究中的“游戏规则改变者”。其主要优势在于减轻了编程负担,大幅提高了研究人员的生产力与效率。消除编程的复杂性,使社会科学研究者能够将更多精力投入到研究的概念化阶段,这可能有助于社会理论的发展。此外,生成式人工智能的变革性影响重塑了研究范式,让计算工具能惠及更广泛的学者群体,进而催生出更多社会科学领域的创新与发现。特别值得一提的是,这些智能工具在教学中的作用也十分显著,生成式人工智能能够为计算学习社会科学课程的学生提供代码注释和解释,显著提高编程和数据分析的学习效果。

然而,这一进展并非毫无前提与风险。尽管有观点认为编程不应构成重大障碍,但学者仍需具备评估代码及其输出质量的基本能力。例如,在为

Swin Transformer 等模型生成代码时,理解深度学习的基本原理至关重要。幸运的是,一门基础的计算社会科学课程通常足以帮助研究者克服这一学习难题。更为紧迫的问题是评估输出质量,特别是在使用诸如 GPT - 4 这样的模型进行视觉和高级数据分析任务时。对于处理大规模图像数据集的研究人员而言,建立系统性验证输出结果的机制很有必要,这一领域本身也值得进一步开展学术探索。

尽管文中提及的一些工具是专有的,但预算限制不应阻碍学术研究。具备一定编程知识的研究者可选择开源替代方案,如 Meta 的 LLaMa 系列,这些模型虽需要一定计算资源在本地执行,但能显著降低成本。对于算力不足的研究者而言,Google Colab 等免费云平台也提供了可行的替代方案,进一步降低了计算社会科学研究的门槛,拓宽了社会科学研究的的可能性。

参考文献:

- Bail, C. 2024, "Can Generative AI Improve Social Science?" *Proceedings of the National Academy of Sciences* 121(21): e2314021121.
- Burke, M. et al. 2021, "Using Satellite Imagery to Understand and Promote Sustainable Development." *Science* 371(6535): eabe8628.
- Davidson, T. 2024, "Start Generating: Harnessing Generative Artificial Intelligence for Sociological Research." *Socius* 10: 23780231241259651.
- Edelmann, A. et al. 2020, "Computational Social Science and Sociology." *Annual Review of Sociology* 46(1): 61 - 81.
- Fan, X. & Y. Zhang 2022, "'Just a Virus' or Politicized Virus? Global Media Reporting of China on COVID - 19." *Chinese Sociological Review* 1: 1 - 28.
- Gilardi, F. et al. 2023, "ChatGPT Outperforms Crowd Workers for Text-Annotation Tasks." *Proceedings of the National Academy of Sciences* 120(30): e2305016120.
- Jean, N. et al. 2016, "Combining Satellite Imagery." *Science* 353(6301): 790 - 794.
- Lazer, D. et al. 2009, "Computational Social Science." *Science* 323(5915): 721 - 723.
- Lazer, D. et al. 2020, "Computational Social Science: Obstacles and Opportunities." *Science* 369(6507): 1060 - 1662.
- Ledford, H. 2020, "How Facebook, Twitter and Other Data Troves Are Revolutionizing Social Science." *Nature* 582(7812): 328 - 330.
- Liu, Z. et al. 2021, "Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows." arXiv:

2103. 14030 [Cs].

- Maher, T. et al. 2024, “The Racial Limits of Disruption: How Race and Tactics Influence Social Movement Organization Testimony before Congress, 1960 – 1995.” *Social Forces* 073.
- Ollion, É. et al. 2024, “The Dangers of Using Proprietary LLMs for Research.” *Nature Machine Intelligence* 6 (1): 4 – 5.
- Seguin, C. et al. 2023, “A Seat at the Table: A New Data Set of Social Movement Organization Representation Before Congress During the Twentieth Century.” *Socius* 9: 23780231221144598.
- Wang, B. et al. 2023, “Towards Understanding Chain-of-Thought Prompting: An Empirical Study of What Matters.”
- Wankmüller, S. 2022, “Introduction to Neural Transfer Learning With Transformers for Social Science Text Analysis.” *Sociological Methods & Research* 00491241221134527.
- Won, D. et al. 2017, “Protest Activity Detection and Perceived Violence Estimation from Social Media Images.” doi:10.1145/3123266.3123282.
- Wu, X. et al. 2024, *Computational Social Science: Application in China Studies*. [S. l.]: Taylor & Francis.
- Yao, S. et al. 2023, “Tree of Thoughts: Deliberate Problem Solving with Large Language Models.”
- Zhang, Y. et al. 2023, “Sinophobia Was Popular in Chinese Language Communities on Twitter During the Early COVID – 19 Pandemic.” *Humanities and Social Sciences Communications* 10(1): 1 – 12.

责任编辑:张继军、赵海峰