

文学计算批评的工具运用与文学研究的范式转型

聂珍钊

(广东外语外贸大学 英语语言文化学院, 广东 广州 510420)

摘要 AI 计算工具在文学研究领域的运用, 给文学创作与批评带来全新变化: 文学分析从文本细读转向图谱解析, 为文学研究提供新的视角和证据, 促进文学与其他学科的融合; 文学批评从主观判断向客观计算的范式转型, 推动文学理解从阐释循环到证据驱动的范式革新, 进而建构全新的文学计算分析批评; 在 AI 文学生成过程中, 创作文学的作者被生成文学的 AI 置换, 人工智能挑战作者的主体地位, 传统文学作品形态向多模态转型; 文学研究方法论在研究范畴、分析维度和知识生产模式上实现三重突破。AI 文学的出现及 AI 工具在文学研究中的深度介入, 将推动以计算分析为基础的文学理论的重构, 助力人类摆脱既有理论体系束缚, 站在科技拓展的认知前沿, 反思并重构文学批评的理论基石, 建立融合智能计算与人文精神的全新理论体系。

关键词 计算批评; 人工智能; AI 文学; 范式转型

DOI: 10.19992/j.cnki.1000-2456.2025.04.006

基金项目 国家社会科学基金重点项目“文学伦理学批评跨学科话语体系建构研究”(21AWW001)

在人工智能(Artificial Intelligence, 简称 AI)推动下, 文学创作与批评理论将发生颠覆性改变, 实现思想观念、研究方法与话语体系的全新建构。AI 技术工具驱动的文学批评新范式, 引发了文学阅读、理解、认知与批评的革命, 具体体现在方法论革命与可视化分析工具的运用、文学观念更新与文学本体论的数字化重构、话语体系重建与文本价值判断的计算转型。从当前发展态势看, AI 技术推动的文学批评理论新范式, 将带来思想观念的更新、计算工具的运用以及计算话语体系的建构, 并导致文学批评理论研究出现三大变化: 一是 AI 技术推动文学的数据化, 文学作品从印刷文本转变为编码数据; 二是由于计算工具和大语言模型的运用, 文学作品的分析从读文转向读图, 即从对文学文本和文献的阅读转向对图表和数据的阅读; 三是从传统的人的主观评价判断转向 AI 客观的计算分析。

一、AI 计算工具运用: 从读文到读图

AI 计算分析批评的基本特点是借助 AI 计算工具阅读文学文本并对其进行批评。AI 计算工具在这里指的是用于文学研究领域的工具, 即利用人工智能技术进行文学计算的工具体, 如专门用于分析海量文本数据和模仿人类创作的人工智能软件。AI 计算工具的功能强大, 能自动生成诗歌或小说文本, 模仿特定作家的风格(如鲁迅的讽刺、马克·吐温的幽默), 对文学作品进行分析, 统计词频、提取主题、分析文本风格与文学文本的心理和情感以及建构文学文本的语言模型等。AI 计算工具种类繁多, 如机器学习框架(如 TensorFlow、PyTorch)、自然语言处理工具、计算机视觉工具、强化学习平台、专门的人工智能芯片、深度学习框架等都可以被视为 AI 计算工具。AI 计算工具可应用于数据清洗与预处理、数据挖掘与特征提取、数据分析与可视化呈现、图像和语音识别、自然语言处理等领域。

大语言模型、文本挖掘算法尤其是思维导图(Mind Map)等工具的介入, 推动了文学分析从传统的文本细读范式向图谱解析范式转型。尽管 AI 计算在原理上同人类认知有某种类似, 但 AI 计算工具对文学文本的解读与人类不同。如对符号的读取和理解, AI 需要经过比人的认知思维过程更复杂的转化。人类

通过文本符号对符号代表的意义进行视觉读取,在语义与情感层面理解文本,而 AI 则借助符号识别系统或文本读取系统,通过算法和模型模拟人类智能,对文本数据进行精确计算,从而实现对文本的认知、学习、理解、分析以及评价。

在计算批评实践中, AI 需要先借助阅读工具将由文字构成的文本转化为结构化数据,再通过建模实现对文本的理解和分析。与 AI 计算分析批评不同,传统文学批评作为基于人类认知机制的实践,其特征涵盖四个维度:认知思维过程(如直觉与辩证思维)、知识存储形式(个人/文化记忆)、表达媒介(语言符号系统)以及社会实践属性(学术共同体互动),其核心在于通过价值判断实现文本意义的人文性生成。无论是中国的“文以载道”,还是西方的“寓教于乐”,都体现了文学认知世界、建构意义的媒介功能。

在阅读过程中,人首先需要借助视觉器官阅读文本并将符号转化为抽象意义。当人的视网膜捕捉到文本符号,这只是对文本的感知,是文本认知的开始。接着,人通过视神经将文本符号转换成电信号传输至人的大脑并在认知神经中枢进行编码,才能获得认知的结果,形成有意义的思想。在这个认知思维过程中,无论编码还是解码,都是根据伦理、政治、文化等标准进行逻辑推理的。从性质上看,这些标准都是不可或缺的伦理标准。离开了这些标准,人的大脑里只能留下一大堆杂乱无章的混乱信息,无法形成清晰的有意义的思想。最后,思想被存储在大脑里,形成固态文本即脑文本,至此阅读文本的认知过程才最终完成。传统批评需要经历从文本符号到脑文本的认知转换过程,而脑文本的呈现,则需要利用人的发音器官将其转换成声音形态即语言,或者利用人的手这个运动器官借助纸张等物质载体将其转换成书写文本。因此,语言和书写文本是表现脑文本的两种主要形式,也可将它们看成表达脑文本的两种方法。

与传统的文学批评不同, AI 借助工具分析和处理文本数据。在进行计算分析批评时, AI 首先将书写符号构成的文本转化为可供计算的数据,然后通过 AI 工具进行分析,分析的关键是建模。在数字人文研究中,模型在具体的数据和抽象理论之间发挥着中介作用,“模型成为理论化的雏形,作为从‘数据资料’到‘理论抽象’的中间环节,被人文学者生产了出来”^①,大数据技术可以处理庞大的文本所形成的数据库,发现人类难以察觉的模式或规律。 AI 计算工具在文学批评中的运用主要表现为对文本的计算以及对计算结果进行阐释和评价,并将计算结果以可视化图表+文本的形式呈现出来。

传统批评注重对文本中书写符号本意及寓意的解读,而在 AI 技术支持下,构成文本的书写符号被分解、标引为一个个数据,并转化为可视化图表,如情感分析生成的情感波动曲线、主题模型输出的语义网络图、风格分析呈现的句法结构矩阵等。借助计算分析工具,文学文本被分解成数据并以图表形式展现出来,这就要求批评者掌握分析数据和图表的能力,比如从色彩、形状、布局等视觉元素中解码文学意义的读图能力。通过读图,如词频分布热力图,能够直观发现《百年孤独》中“孤独”主题的叙事密度与节奏变化。通过文本挖掘算法,解读莎士比亚的戏剧可通过对剧中台词的词性、词频等数据进行提取分析,生成人物对话的关系网络图,从中可以清晰地发现不同角色之间对话的频繁程度和关系亲疏,这为研究人物关系和戏剧冲突提供了全新视角,也可以直观展现读图在文学分析中的优势。在《远读》(*Distant Reading*)一书中,弗朗哥·莫雷蒂(Franco Moretti)通过绘制《哈姆雷特》人物关系网络图的方法分析人物的重要性,图中哈姆雷特处在中心位置。他说:“如果没有哈姆雷特这个角色,我们绝不会想到去讨论《哈姆雷特》这部作品。”^②他还通过网络图发现,同哈姆雷特和克劳狄斯相比,“霍雷肖恰恰相反:他所在的网络部分聚类程度极低,要是没有他,这部分网络就会瓦解”^③。从中可以看出,读图不仅为理解文学文本提供了一种可视化方法,而且也为挖掘文学作品的主题和价值、理解人物和情节提供了新的思路。

传统文学批评和研究强调对文学文本的细读和深入剖析,通过文字来感受作者的情感、理解作品的主题和意义。但随着数字化时代的来临,由于计算工具和大语言模型的运用,文学作品的分析逐渐转向读图,即通过对图表、数据可视化等形式的阅读来把握文学作品的内在规律和外在表现。图表和数据能够直观地展示文学作品的统计特征、结构关系、情感分布等信息,为文学批评提供新的视角和证据。通过读图,

① 赵薇:《作为计算批评的数字人文》,《中国文学批评》2022年第2期。

② Franco Moretti, “Network Theory, Plot Analysis,” in Franco Moretti, *Distant Reading*, London & New York: Verso, 2013, p. 220.

③ Franco Moretti, “Network Theory, Plot Analysis,” in Franco Moretti, *Distant Reading*, p. 227.

研究者可以更加客观地分析文学作品的语言风格、叙事结构、人物关系等,减少主观判断的偏差。“文学阐释需要依赖阐释者的主观判断,涉及阐释主体的经验、学识、判断力,而数字人文则通过量化分析和可视化表达,为文学研究提供了更为客观和理性的依据。”^①从读文到读图的转变促进了文学与计算机科学、数据科学等其他学科的融合,为文学研究提供新的方法和工具。

从读文到读图的转向,不仅是研究方法的革新,更是学术观念的范式转换。莫雷蒂在《图表、地图、谱系树:文学史的抽象模型》(*Graphs, Maps, Trees: Abstract Models for Literary History*)一书中,以书名所指的三种可视化模型(图表、地图、谱系树)为工具,对文学作品展开了方法论层面的实践研究。他认为,这种可视化研究将促使文学研究者重构工作路径:“我们得以拓宽文学研究的疆域,进而以时间、空间、形态等新维度的分类体系,取代传统研究中低效的二元对立(诸如高雅与低俗、经典与档案、民族文学分野……)。”^②

同文本阅读相比,可视化的读图不仅更容易理解文本、人物、风格、主题,而且更容易对文本内容、人物关系、故事情节、主题思想、文本价值进行总结归纳。研究者现在倾向于用图表、思维导图、知识树等对文本进行归纳、总结和评价,用新的形态、术语和表达取代旧有的形态、术语和表达,用科学的方法取代传统的人文方法,让人耳目一新。新技术、新工具在文学研究中的运用,可以发现传统方法仅凭人力无法发现的规律、现象和结论,这不仅提高了研究的效率,更带来了文学观念的革新。传统文学批评依赖批评家在文本阅读中形成的主观判断和经验积累,主观性和不确定性较大。在数字化时代,文学分析逐渐转向客观的计算分析,通过算法模型对文学作品进行量化处理,将文本转化为更直观的图表。因此,像诊断病情的医生一样读图和数据列表,将成为未来读者或批评家理解、分析和评价文学的基本方法和技能。

二、范式转型:从文本阐释到数据计算

计算工具的大量使用,将促使文学批评从主观判断向科学计算的范式转型。AI技术给文学批评带来的根本性变化,是要求建构文学批评的计算模型,将文学文本的主观阐释转变为文本数据的科学计算。在数字革命的背景下将文学与科学进行结合,不仅是传统文科教育实现转型升级、提高学科地位的新思路,而且为文学研究提供了新的范式:“数字人文的勃兴,是人文学者对被边缘化学科自我救赎,而其自救所用的数据计算方法,则为探讨文学的科学性提供了可能。”^③

传统的文学研究强调读者或批评家的个人判断能力,批评者通过掌握的理论方法阅读和理解文学文本,进而做出主观评价。所谓“一千个读者就有一千个哈姆雷特”,表明传统文学批评既具有个人主观性,也具有任意性。从本质上看,这不仅说明传统文学批评具有个性化特点,而且说明这种批评缺乏公认的评价标准。而科学的评价必然是有标准的。

传统批评依赖批评家的主观判断,而计算分析则依赖概率统计、模式识别以及客观数据的计算。例如,利用机器学习评估诗歌韵律时,算法会计算音节、重音、押韵等特征的匹配度,生成量化评分。这种转变并非否定人文阐释,而是将主观经验与客观计算相结合,形成“人机协同”的评价。计算批评,并非强调如何分析和表达文学的审美,而是致力于建构符合计算逻辑的新型批评话语体系。例如在解释小说叙事结构时,既能分析情节安排的艺术性,又能呈现算法识别的叙事弧线的数学特征。再如对唐代诗歌风格进行研究时,可以运用机器学习算法对大量唐诗的文本数据进行处理。算法从诗歌的词汇运用、句式结构、韵律特点等多方面提取特征,通过概率统计分析不同诗人诗歌风格的差异。比如对李白和杜甫诗歌进行对比分析时,计算模型能够敏锐捕捉到李白诗歌用词的豪放洒脱、韵律的灵动多变和杜甫诗歌用词的沉郁顿挫及格律的严谨规范等,提供客观量化的数据,从而改变仅靠主观感受评价诗歌风格的方法。

计算分析批评模型是运用计算方法和技术分析文学作品的理论框架和研究工具,通过对大规模文学文本数据的挖掘、分析和建模,揭示文学作品的内在结构、主题演变、风格特征以及作者的创作模式等,是未来文学研究基本的科学方法。计算批评模型即计算分析批评的方法,可用于文学的全面分析研究。例

① 卓今:《数字人文的文学阐释:情感计算、数字信任及基本实现手段》,《湖南师范大学社会科学学报》2025年第2期。

② Franco Moretti, *Graphs, Maps, Trees: Abstract Models for Literary History*, London & New York: Verso, 2007, p. 91.

③ 彭青龙:《数字时代文学的科学性与人文性审视》,《探索与争鸣》2022年第9期。

如,利用潜在狄利克雷分配(LDA)等主题模型,能够自动发现文本中潜在的主题结构;利用机器学习算法,如支持向量机、朴素贝叶斯分类算法等对文学文本进行情感分类训练,不仅能学习到文本特征与情感类别之间的映射关系,还能根据人物的语言、行为以及周围环境的描述,更准确地判断其情感倾向。计算批评模型能从大规模文本数据中发现传统研究方法难以察觉的模式和规律,如在分析小说人物的对话或内心独白时,机器学习的情感分析模型可更准确地判断人物的心理变化轨迹和情感状态,从而帮助人们更好地理解人物的性格和心理变化。

在 AI 大模型里,情感是可以计算的。情感计算(Affective Computing)是通过算法量化分析文本、语音或图像中情感状态的技术。在文学分析中,情感计算可帮助我们理解角色在不同场景中的情绪波动。例如,当大语言模型分析《哈姆雷特》时,其生成的“人物关系图谱”不仅能呈现权力斗争的结构,更能通过情感分析算法,将“生存还是毁灭”(To be, or not to be)独白中的矛盾心理与当代抑郁症状的语言模式进行跨时空对比,进而理解哈姆雷特在两难选择中的复杂情感。对托尔斯泰《战争与和平》的研究,可以运用 LDA 主题模型进行分析,从海量文本中挖掘出战争、爱情、贵族生活、历史变革等多个主题,并清晰呈现这些主题在章节中的分布与演变。通过对人物对话和内心独白的情感分析,机器学习模型能准确捕捉到人物在战争背景下复杂的情感变化,为深入发掘、理解小说的多主题结构和人物塑造提供科学且深入的视角。

引入计算分析方法,能够在物质形态的科学观念基础上构建文学研究的客观解释体系,将主观阐释转化为可计算的神经认知活动分析,将文化相对主义的价值判断转化为跨文化的生理—心理反应式研究,最终实现文学理解从阐释循环到证据驱动的范式革新,进而建构全新的文学计算分析批评。

三、AI 文学的生成:人工智能对作者主体的挑战

AI 引领传统文学创作与研究向人机协同创作的新形态转型,催生出由算法生成的 AI 文学或者 AIGC 文学。这给传统的文学概念带来了巨大冲击,甚至在一定程度上动摇了文学作为人类创造物的属性:“从本质上说, AI 文学是算法文学。算法的品质决定 AI 文学的品质。”^①AIGC 大模型作为生成文学文本的 AI 工具,通过模拟多种文学风格生成算法文本,消解了传统文学创作中作者的意图,否定了作者决定文学文本的唯一权力。在 AI 文学生成过程中,创作文学的作者被生成文学的 AI 置换或部分替换, AI 成为文学生成的主体或“拟主体”。在传统观念里,没有作家的参与,文学创作是不可想象的,而借助大语言模型的算法机制, AIGC 也能生成只有作家才能创作的文学文本,且文本的品质并不逊色。这表明机器通过学习海量文学数据,找到了理解语言的密码和组织文本的方法,能够建构超越人类直觉的文本模式,既能生成具有共性特点的文学文本,也能生成具有个性特点的文学文本。

当前人工智能技术的飞速发展在一定程度上模糊了人与机器的界限,“人工智能使得机器从机械物跃升为智能人,人与机器的边界被彻底打破,机器似乎更像人,而人似乎更像机器”^②。文学创作也是如此,作家同人工智能的边界正在被打破。在传统观念中,文学由作家创作,没有作家就没有文学作品。但 AIGC 用其生成的文学文本证明,文学创作机制正在被改变,文学基本原理正在被重构,作家的创作实践可能被 AIGC 取代。在由算法生成文学文本的过程中,作家主体地位遭到严重挑战,随着算法的迭代升级, AI 正趋近于文学创作的主体, AI 文学正在成为具有迭代特征的文学类型。

AI 文学的生成虽有人类作家参与,但它并非传统意义上由人类作家创作的文学文本。在 AIGC 的文学生成过程中,作家与文本关系的传统认知被颠覆,作家作为创作主体的价值被消解,作家中心主义的观点在 AI 算法的冲击下面临失守之势, AIGC 正在逐步变成 AI 文学的主体并决定文学的生成。

长期以来,人们认为文学作品的意义和价值主要取决于作家的生平、思想、创作意图等因素,文学批评的重点需要关注对作家创作背景和心里的分析。20 世纪,作家中心主义遭遇文本中心主义的挑战。形式主义、新批评和后结构主义的理论基点从作家转向文本,主张文学研究的核心对象应是文本本身,而非作

^① 曾巍:《AI 文学的批评:路径、范式与重点》,《云南社会科学》2024 年第 5 期。

^② Gexin Yang, “From Ethical Selection to Scientific Selection,” *Interdisciplinary Studies of Literature*, vol. 6, no. 3, 2022, pp. 416-425.

者、读者或社会历史背景,文学意义应从文本的语言结构、修辞技巧和内在逻辑中寻找,而非依赖作者意图、读者反应或社会语境等外部因素,推动文学研究逐渐转向文本、读者和文化分析。

21 世纪,尤其是以 ChatGPT、文心一言为代表的 AI 文本生成,将文本中心主义推向顶峰,文本的作者似乎已无关紧要。AIGC 通过算法生成文学的机制,与作家通过思维创作的机制截然不同,彻底颠覆了基于作家中心主义建立的文学创作传统。从 ChatGPT 大语言模型到 Sora 世界模型,AI 不仅重构了文学创作的物质基础,更在认识论层面推动着文学批评话语体系的范式转换。这种转变已突破单纯工具革新的范畴,演变为涉及文学本体论、认识论与方法论的系统性变革。在创作实践层面,AI 生成的文本呈现出传统文学未曾触及的“非人类美学”特质。谷歌 DeepMind 开发的“故事编织者”系统能够基于用户输入的提示词实时生成非线性叙事,其作品结构突破线性逻辑束缚,形成类似博尔赫斯“小径分岔的花园”式的叙事迷宫。这种由算法驱动的“可能性写作”,正在催生新型文学体裁,迫使理论界重新审视罗兰·巴特提出的“作者已死”的命题。广东外语外贸大学未来人文实验室开发的“白云深处”(DeepCloud)系统,借助 AI 克隆技术让那些去世多年的伟大天才作家“复活”,重新开始新时代的创作,同时还能训练集无数天才作家“基因”于一身的数字作家,生成满足现实需要并体现时代特色的新形态文学作品。AI 生成文学文本的能力在科幻文学领域表现得尤为突出。日本作家 Rie Kudan 使用 ChatGPT 辅助创作的小说《东京共鸣塔》荣膺日本文学界最高奖“芥川奖”。AI 激发了 Kudan 的灵感,使她得以更好地发挥自己的潜力,成为她创作过程中得力的“伙伴”。清华大学新闻传播学院教授沈阳借助 AI 平台创作的科幻小说《机忆之地》,在第五届江苏省青年科幻作品大赛中获得二等奖,大多数评委都未从中发现 AI 创作的痕迹。在小说的生成过程中,具备多领域知识、多学科视角和较强逻辑推理能力的 AI,不仅可以提供科幻写作所需的庞大知识数据,而且在情节设计、心理描写、细节丰富性等方面表现出巨大潜力,为科幻文学创作提供了强大的技术支持。

在 AIGC 生成文学的机制中,能看到文学创作在完全智能化之前正在经历“人机协同创作”阶段,文学研究也从文本阐释转向“生成过程追踪”,如神经网络的可解释性分析。例如,AI 已从辅助工具如语法检查发展为创作伙伴,如 GPT-4 生成完整小说章节。这意味着文学研究不再仅仅关注文学作品本身的内容和意义,而是深入探究文学作品如何通过 AI 技术生成的过程,包括其中涉及的算法、模型以及各种技术参数等对作品生成的影响。

AIGC 生成文学从“人机协同”到“完全生成”,是 AI 文学从人机协同到完全智能化的演进过程。当前 AIGC 的文学生成仍需要人的参与协同,即需要人提供主题、形式、内容、风格、修改等方面的要求以及基本资源。在 2020 年之前,AI 作为基础工具,主要任务并非文本生成,而是承担语法校对、风格建议(如 Grammarly)、简单续写(如早期 GPT-2)等功能。作为人类作者创作的辅助工具,离开人的参与,GPT 很难创作出合格作品。但 GPT-2 的发布为后续的 GPT-3(2020 年 6 月)及 ChatGPT(2022 年 11 月)奠定了基础,推动了生成式 AI 的发展。人类通过提示词(prompt)主导生成,形成人类设定主题和框架,AI 填充内容,人类引导修改润色的“闭环能力”,即 AIGC 的“创作—评估—优化”能力。当前的 GPT-4、Claude、文心一言、DeepSeek 等模型,不仅能生成完整段落、章节甚至整部书籍,还能实时响应修改需求(如 Sudowrite),支持多轮迭代优化,生成过程呈现“对话式”特征。随着 AI 更新迭代加速,人机协同出现人的干预越来越少而 GPT 文本生成智慧化越来越高的趋势。至此,AIGC 能够实现从独立生成作品到对作品进行评估,再根据评估结果进行优化的全流程,无需人工过多干预,从而实现自我迭代和不断提升。

AI 作为新的文学生成的智能体,表现出巨大的文学生成能力,尤其是文学的创新能力。18 世纪英国启蒙主义诗人扬格曾经批评英国文学,认为是盲目崇拜和仿效古典名著导致作家无法创作出具有独特个性的伟大作品^①。扬格指出了迄今为止文学创作的一个残酷事实,文学独创性的缺失源于作家对作家的模仿。然而 AIGC 能通过自主学习,集古今中外无数作家的知识于一身,也能通过算法避免拙劣的模仿,因此能够生成具有独创性的作品,尤其是生成多模态文学作品。同作家的模仿相比,AIGC 通过建模将模仿变为创新,尤其是通过作家不具备的能力生成多模态文学,引领传统文学作品的新形态转型。

当前,AIGC 用于文学创作的主要目的是生成文学文本,而随着更新迭代,AI 文学在形式上与传统文学最大的不同将是多模态呈现,“作为新技术的产物,人工智能文学自身最合适、最自然的存在形态必然是

^① 参见爱德华·扬格:《试论独创性作品》,袁可嘉译,北京:人民文学出版社,1963 年,第 8-10 页。

数字化的,而且必然是多媒介和多模态的”^①。这种 AI 文学运用自然语言处理(NLP),计算机视觉、音频合成等融合技术,能够生成包含文本、图像、音频、视频交互的多模态文学作品。当前 AI 已能实现文本—图像—音乐的跨模态映射,如 OpenAI 推出的新一代图像生成模型 Stable Diffusion 和 DALL·E 3,都能很好地理解图像细节和微妙之处,轻松将创意转化为精美图像,或根据小说情节自动生成配乐与插画。与单一模态的文本相比,多模态文学作品拥有更加复杂的意义潜力,“多模态信息蕴含的知识不仅比单模态更为丰富,而且能衍生出隐性知识”^②。它不是单一模态文本意义的简单相加,而是通过不同模态之间的动态交互协同构建出新的意义。

AI 视频生成是人工智能在视听领域的前沿应用,其核心在于通过算法模拟人类创造力,将概念转化为动态影像。这项技术正在重塑影视制作、广告营销、虚拟现实等多个行业,其发展历程折射出技术与艺术的深度融合。长期以来,文本既是文学载体,也是文学的形式,而将 AI 视频生成应用于文学生成,不仅从根本上改变了文学的载体,而且改变了文学的形式,体现了 AI 文学的多模态特点。

随着 AI 文学的出现,人类的作者身份遭到了挑战,尽管这并不意味着人类将被驱逐出文学的领域,但它表明 AI 文学的生成必将会在文学创作中占据一席之地。AIGC 生成文学在原理上同作者创作文学是类似的,同样需要学习、搜集创作素材、模仿以及需要呈现。经过深度学习和预训练,AIGC 能够掌握许多作家所掌握的艺术技巧,并能熟练地运用这些技巧生成文学文本。我们会发现一种趋势,人类的文学书写会越来越离不开 AIGC 的协同或 AIGC 的引导。当然,未来的文学书写还在发展之中,但 AIGC 作为文学生成的智能体,将在文学创作中发挥越来越大的作用,这已是不争的事实,问题仅在于我们是否承认它、接受它。

四、计算分析的运用与批评体系的建构

文学研究方法向计算分析转型,不仅改变了文学研究的认知框架,更在方法论层面实现了三重突破:在研究范畴上,突破了传统细读法受限于人类阅读能力的样本规模,使跨时代、跨语种、跨文类的全景式分析成为可能;在分析维度上,通过主题建模、情感分析、风格计量等技术,将原本依赖主观感受的文学要素特点转化为可量化、可验证的数据指标;在知识生产模式上,构建起人机协同的阐释系统,其中神经网络模型负责发现文本潜藏的关联模式,研究者则致力于解读数据背后的文学逻辑。

数字人文驱动下文学研究方法论的三重突破,正通过系统性重构研究路径与认知框架,推动文学研究从主观经验判断向客观数据计算的范式转变。有学者将这一过程分为“数学化”“量化—统计化”和“数字化—数据化—向量化”三个阶段,并指出经过这三个阶段的发展,文学研究必将同计算分析融合:“人工智能技术提出了一个新的诠释学的使命。这一新的诠释学使命可以被命名为走向文论的‘算法阐释’。”^③这里的算法阐释指的就是计算分析,它不仅意味着计算工具的运用,更标志着文学批评体系在认识论层面的范式革命,其价值核心在于构建“数据驱动—算法分析—文学批评”的 AI 文学文本研究新范式。

人工智能对文学文本语义的解读和判断同人类大脑有着本质的不同,它反映的是机器智能同生物智能在认知原理、知识构建和价值判断层面的根本区别。“语义解读”是对语言所表达的意义进行分析和理解,是文学批评的基础。传统文学批评依靠读者或批评家的主观解读,结合文本的语言、结构、历史背景等因素,揭示文本的主题、象征意义和社会内涵。由于人脑能力的有限与认知偏差,读者或批评家对文本的误读、误解、误释不可避免,这就决定了人脑对文本的理解是不完全的、是可能存在谬误的。在文学研究中,AI 的介入改变了语义解读的方法。通过自然语言处理和机器学习技术,AI 可以分析大量文本数据并提供客观的语义分析。尤其是通过主题建模、情感分析、风格计量等技术,AI 可以将原本依赖于主观感受的文学文本阐释转化为可量化、可验证的数据计算,通过计算对文本给予客观和科学的解释。这种技术被应用于文学批评中,可以帮助批评家发现文本中的隐藏主题、情感倾向和风格特征。AI 如同高倍显微镜,在文学批评中可以让批评家看见文本肌理中的 DNA 序列,解读文学文本的底层密码。

① 李国成:《人工智能文学及其对现代文学观念的挑战》,《中国社会科学》2024 年第 7 期。

② 苏祺、位通:《数字人文研究的多模态转向》,《山东社会科学》2024 年第 10 期。

③ 曾军:《算法阐释:人工智能时代的文论问题》,《华中师范大学学报(人文社会科学版)》2023 年第 5 期。

同传统研究相比,计算分析能够逐渐将文学研究的时空扩展至传统细读法难以企及的微观和宏观层面。当主题建模算法处理十海量级的维多利亚时期小说时,研究者得以揭示工业革命背景下“机械隐喻”的跨文本扩散规律。词嵌入技术对《四库全书》的语义拓扑分析,可以将词语之间的语义关联以拓扑结构的形式呈现出来,比如可以发现“仁义”概念在不同文本中的连接密度的差异。从“抽样阐释”到“全量分析”的转变,打破了新批评派将文本视作封闭容器的预设,不再将文学现象看成孤立存在,而是把它置于更大的社会文化网络背景中,考查文学现象与社会文化网络的动态联系。斯坦福文学实验室(Stanford Literary Lab)运用词向量模型对 18 世纪情感小说进行语义拓扑分析,揭示了传统细读法难以察觉的叙事模式演变规律;通过地理信息系统(GIS)建模的方式对 19 世纪全球小说市场进行研究,更清晰地展现出不同地区小说市场的特点和变化,能够发现畅销书出现在巴黎左岸书店和伦敦同类书店的时间差异,将文学地理学的研究精度从国家尺度提升至城市街区的维度,实现了文化传播机制的空间计量分析。清华大学开发的“文心”系统通过对《全唐诗》的格律建模,发现了初唐到盛唐时期诗歌形式演化的量化拐点。

在分析维度的量化转型中,算法将文学要素转化为可计算对象,可建立连接人文感知与数学模型的解释桥梁。情感分析算法利用 BERT 模型的情感极性检测功能,可以对意识流文本进行分析,从而确定其中的情感倾向是积极、消极还是中性等,并描绘出该文本中情绪强度随着文本内容变化而产生的起伏,并以曲线的形式直观呈现出来。例如对伍尔夫的小说《到灯塔去》进行分析,可以将其中隐含的焦虑感转化为能够进行验证的时间序列数据。具体而言,就是从不同时间段内人物的言行、心理活动内容中提取出与焦虑感相关的特征,然后将这些特征在不同的时间点上进行量化,形成一系列的数据点,这样就可以对小说中焦虑感的变化情况进行分析和考察。再如风格计量学运用词频分布和句法复杂度计算这两种方法,能够为《红楼梦》后四十回的著作权争议提供统计学证据。这种量化转型并非消解文学性,而是通过算法构建“文学特征空间”,把原本难以明确表述的审美感受转化为如词频分布、句法复杂度计算等可以进行量化和交叉验证的维度,从而对审美体验进行具体的计算分析。例如分析一首诗歌的审美体验,可以将诗歌中的隐喻编码为概念映射网络,即把隐喻所涉及的不同概念之间的对应关系以网络的形式呈现出来。哈佛隐喻识别项目就是通过将诗歌隐喻编码为概念映射网络,为传统批评中“只可意会”的修辞效果找到了可计算的解释方法。

计算分析体现的是强大的科学力量,不仅没有消解读者和批评家的作用,反而突出了人的作用,强化了人机协同的机制,催生了“增强型阐释”范式。传统的文本细读分析依赖批评主体在阅读过程中的经验,而算法的分析不仅要依赖人工智能的经验,更要依赖“人机融合的经验”,即“人与算法软件在虚实世界中沉浸式的具身交互共融共生”^①。因此,人机协同也不是简单的工具辅助,而是要在两种阐释方式之间切换比较,形成“算法生成假设—学者验证解释”的认知闭环模式。例如对魔幻现实主义文学的跨文化传播研究,人机协同的运作方式是通过模型先识别超现实元素分布规律,研究者再结合地缘政治进行文化解码,最终构建解释模型。算法的参与非但不会损伤作品的文学性,甚至还可以挖掘出隐藏在文本更深层的审美潜力,这需要人的参与才能实现。“人机协同”不是简单地将文学作品抽象为数据来进行分析,它也涉及将两种阐释方式进行混合与迭代,对同一对象反复进行混合,“这种方法可以让我们重新认识阅读单一文本的(审美)愉悦,拒绝贬低文本中的无限世界,这个世界等待着我们去探索、解释、再探索”^②。

无论是文本、音频还是视频文本,都可以编码成可计算的数据。通过对文本数据的计算,量化分析技术可揭示文学创作的深层结构特征,从而建立解读文学的科学方法。以英国浪漫主义诗歌的计量研究为例,研究者通过构建自然语言处理模型对拜伦、济慈、华兹华斯、雪莱等诗人的语料库进行主题建模和情感分析,可以发现自然意象的呈现方式具有显著的风格辨识度。具体而言,华兹华斯诗作中“溪流”“牧群”“雏菊”等意象的出现密度与“平静”“安谧”等情感词汇形成强正相关,其文本向量空间呈现出显著的田园叙事聚类特征;而雪莱作品中“风暴”“雷电”“火山”等意象则与“自由”“变革”等政治隐喻词汇高频共现,其语义网络显示出强烈的冲突性修辞模式。这种基于远读范式的风格计量研究,通过词频统计、共现矩阵和

① 蒋济永、王磊:《数字人文与文学批评的互文性表达及 AI 生成式诗学构建》,《东南学术》2025 年第 1 期。

② Moacir P. DE SÁ Pereira, “Mixed Methodological Digital Humanities,” in M. K. Gold and L. F. Klein, eds., *Debates in the Digital Humanities*, Minneapolis: University of Minnesota Press, 2019, pp. 405-412.

主题建模等计算方法,为传统印象式批评提供了可验证的量化证据,标志着文学研究方法正经历从定性阐释向计算驱动的范式转型。这种方法论创新不仅实现了文学批评从定性到定量的范式转换,更重要的是建立了文学现象与历史语境的关联模型,为跨学科人文研究提供了示范。

在数字化技术重构知识生产范式的时代语境下,基于智能分析工具与大语言模型(LLMs)建构的新型文学批评体系,以数据客观性为方法论基石,通过算法运算的精确性和多学科交叉的阐释框架,推动文学研究向科学化、精准化方向转型。在文学研究中,计算工具和大语言模型可对海量文学语料进行清洗、向量化与模式挖掘,深度神经网络不仅能捕捉显性的题材偏好与修辞特征,更能揭示文本中潜在的情感光谱、认知框架与形式演化规律。过去仅凭人力和传统方法无法完成的任务,如今借助AI技术,研究者可快速获取文学作品的相关数据,通过计算分析,准确揭示文学作品的隐含主题、情感倾向、风格特征等深层次信息,实现对文学作品的科学阅读理解。

AI文学的出现及AI工具在文学研究中的深度介入,正在引发一场链式反应般的范式革命,推动文学形式、观念与研究方法的变革,以及以计算分析为基础的文学理论重构。这是AI时代文学创作与研究的必然转向,其速度远超预期,它既是对传统文学研究方法的挑战,又为文学批评和文学理论的进一步发展提供了新的视角。当AI能够模拟人类认知且能超越传统认知框架时,人类就能够跳出既有理论体系,进入更广阔的科学阐释空间。这种突破促使我们站在科技拓展的认知前沿,反思并重构文学批评的理论基石,将可验证的数据思维融入文本解读的方法和过程之中,最终建立起融合智能计算与人文精神的全新理论体系。

责任编辑 曾新

The Application of Computational Tools in Literary Computational Criticism and the Paradigm Shift in Literary Studies

Nie Zhenzhao

(School of English Language and Culture, Guangdong University of Foreign Studies, Guangzhou 510420)

Abstract : The use of AI computational tools in the field of literary research has brought about novel changes to literary creation and criticism; literary analysis has shifted from textual close reading to graphical analysis, providing new perspectives and evidence for literary research and promoting the integration of literature with other disciplines; literary criticism has undergone a paradigm shift from subjective judgment to objective computation, driving a transition in literary understanding from interpretive cycle to evidence-driven innovation, thus constructing a brand new framework of literary computational analysis criticism; in the process of AI-generated literature, the traditional author who creates literature is replaced by AI as the generator of literature, challenging the author's subjective position, and transforming the traditional literary works into the multi-modal; the methodology of literary research realizes a triple breakthrough in research scope, analytical dimension, and knowledge production mode. The emergence of AI literature and the in-depth intervention of AI tools in literary research will promote the reconstruction of literary theory based on computational analysis, help to free human beings from the bondage of the established theoretical system, stand on the cognitive frontier of technological expansion, reflect on and reconstruct the theoretical cornerstone of literary criticism, and ultimately establish a new theoretical system that integrates intelligent computation with humanistic spirit.

Key words : computational criticism; artificial intelligence; AI literature; paradigm shift