复杂系统传播动力学中的大语言模型

Shuyu Jiang[1,2,3], Hao Ren[1,2,3], Yichang Gao[4], Yi-Cheng Zhang[5],Li Qi[6],Dayong Xiao[6],Jie Fan[*7], Rui Tang[*1, 2, 3], Wei Wang[*8]

1. 四川大学网络空间科学与工程学院，中国成都，*610065 2.* 数据保护与智能管理重点实验室（四川大学），教育部，中国成都，*610065 3.* 四川大学网络空间科学研究院，中国成都，*610065 4.* 皇家墨尔本理工大学，*GPO Box 2476*，澳大利亚维多利亚州墨尔本 *3001 5.* 弗里堡大学物理系，*Chemin du Musée 3*，瑞士弗里堡 *1700 6.* 重庆市疾病预防控制中心（重庆市预防医学科学院），中国重庆 *400070 7.* 重庆医药高等专科学校公共卫生与应急管理学院，中国重庆 *401331 8.* 重庆医科大学公共卫生学院，中国重庆，*400016*

arXiv:2602.08085v1 [physics.soc-ph] 8 Feb 2026

摘要

传播动力学是复杂系统物理学和网络科学的核心议题，为理解信息、行为和疾病如何通过系统单元之间的互动进行传播提供了统一框架。在许多传播情境中，传播过程受到多种相互作用因素的影响，例如信息表达模式、文化背景、生活环境、认知偏好和公共政策，而这些因素难以直接纳入经典建模框架。近年来，大语言模型（LLMs）在自然语言理解、推理和生成方面展现出强大能力，能够明确感知传播过程中的语义内容和语境线索，从而支持对不同影响因素的分析。除了作为外部分析工具，大语言模型还可充当嵌入传播系统的交互式智能体，可能影响传播路径和反馈结构。因此，大语言模型在传播动力学中的作用与影响，已成为多个研究领域中一个活跃且迅速发展的研究方向。本综述全面概述了将大语言模型应用于传播动力学研究的最新进展，重点关注两个代表性领域：数字流行病（如错误信息和谣言）以及生物流行病（包括传染病暴发）。我们首先从复杂系统视角考察流行病建模的基础，并讨论基于大语言模型的方法与传统框架之间的关系。随后，我们从三个关键视角系统综述近期研究，即流行病建模、流行病检测与监测，以及流行病预测与管理，以阐明大语言模型如何促进这些领域的发展。最后，我们讨论尚待解决的挑战和潜在研究方向。

关键词：大语言模型、传播动力学、数字流行病、生物流行病、复杂系统、生成式智能体、网络科学、计算社会科学

[1]本论文已被《物理学报告》接受发表。

*通讯作者：241031@cqmpc.edu.cn（Jie Fan），tangrscu@scu.edu.cn（Rui Tang），wwzqbx@hotmail.com（Wei Wang）

目录

1 引言 **3**

2 背景 **7**

2.1 传播动力学导论.7 2.1.1 基本概念.7 2.1.2 传播动力学经典模型.9 2.2 大语言模型导论.10 2.2.1 使大语言模型具备功能的过程.11 2.2.2 大语言模型在流行病相关任务中的一般工作流程.12 2.3 用于验证的经验数据.13

3 用于流行病建模的大语言模型 **15**

3.1 基于大语言模型的数字流行病建模.16 3.1.1 大语言模型作为数字流行病建模的增强工具.17 3.1.2 大语言模型作为数字疫情传播的影响因素.20 3.2 基于大语言模型的生物疫情建模.21 3.2.1 大语言模型直接参与建模过程.21 3.2.2 大语言模型作为辅助工具间接参与流行病建模.22

4 流行病检测与监测：大语言模型增强的传播动力学感知 23

4.1 大语言模型增强的数字流行病感知.23 4.1.1 感知传播内容与语境.24 4.1.2 感知传播网络.26 4.1.3 感知社会认知.26 4.2 大语言模型增强的生物流行病感知.26 4.2.1 早期检测.27 4.2.2 持续疫情监测.28

5 流行病预测与管理：大语言模型增强的传播动态调控 305

.1 大语言模型增强的 hanced digital epidemi传播调控lati

上..... 30 5.1.1 用于预测的大语言模型..... 31 5.1.2 用于管理的大语言模型..... 35 5.2 大语言模型增强的生物疫情调控..... 37

6 结论与展望 **40**

1. 引言

在我们的世界中，各种各样的物理、社会、生物学和技术系统都呈现出复杂结构和交互 [1]。由于这些系统无处不在且十分复杂，它们持续吸引着物理学家、生物学家、生态学家、经济学家和社会科学家的关注 [2]。研究人员感兴趣的不仅是个别系统，也包括一些系统类别；这些系统虽性质各异，却具有相似的结构特征 [3]。这种兴趣源于一项长期挑战：如何准确、全面地描述不同的复杂系统。这一问题长期以来被视为人类面临的重大科学难题之一 [4]。为应对这一挑战，科学家一直致力于重新构建这些系统，至少通过数学模型来概括整个系统集合的基本特征 [1； 4]。事实上，尽管不同的复杂系统表现出多种多样的行为和现象，其底层数学表示通常依赖于相似的建模范式 [5]。在这些范式中，复杂网络尤为突出，是一种特别有效且具有代表性的方法。最简单的形式是将系统的基本单元表示为节点，并用边描述它们之间的互动，从而为分析结构和动力学性质提供简洁的抽象 [6; 7]。

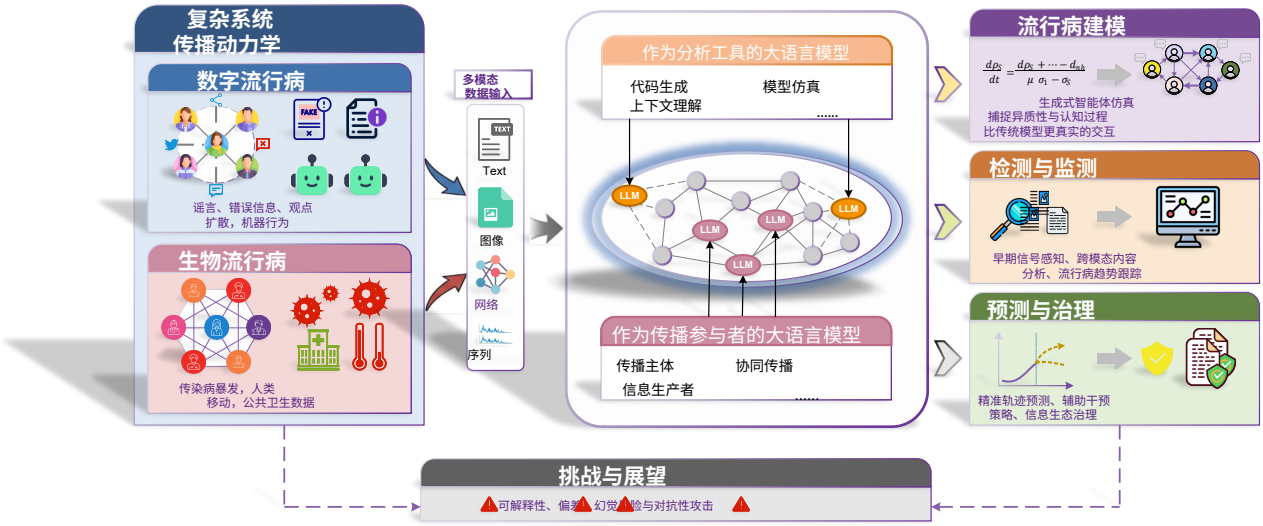


图1：论文结构概览，展示了融合大语言模型的传播动力学分析。左侧展示了两类具有代表性的传播过程：生物疫情和数字流行病。中央部分展示了异构数据源，包括文本、网络和多模态信号，它们为传播过程提供观测数据。大语言模型被描绘为核心组成部分，与这些数据进行交互，既是处理和推理传播信息的分析工具，也是可能参与信息生成与传播的主动实体。右侧展示了大语言模型增强型应用在传播工作流关键阶段的布局，包括建模、检测、监测、预测和治理。底部总结了与大语言模型增强型传播系统相关的代表性挑战和风险。

由于复杂网络是复杂系统的有力表示，这些系统内在的传播现象可以通过相应网络结构上展开的传播动力学进行理论刻画；传染病或信息（尤其是低风险信息）等传播对象被视为在节点之间流动或扩散的“流行病” [8； 9； 10； 11； 12； 13； 14； 15]。更具体地说，不同个体节点之间的互动使传播过程呈现出结构化的时间模式和群体特征。刻画传播动力学不仅是对个体行为的描述，也是理解群体系统如何随时间演变的重要途径。通过分析传播动力学，科学家可以观察群体如何形成共识或产生分歧、网络如何爆发或衰退、风险如何在不同地区和社区之间流动，以及局部扰动如何触发系统性响应。在此基础上，系统地揭示传播过程的潜在模式和关键机制，为预测其未来轨迹、缓解不良传播、制定有效的管理策略等奠定基础 [16]。

经典传播动力学研究通常借鉴统计物理学、网络科学、应用数学和流行病学的理论，对复杂传播过程进行简化和抽象，进而构建从简单的舱室模型（如易感者–感染者–易感者（SIS）模型和易感者–感染者–康复者（SIR）模型）到大规模基于主体的模型（ABM）的各类模型；后者可模拟个体在城市内部或跨国之间的移动 [17; 18; 7]。这些模型能够从理论上解释基本机制如何影响重要特征，在揭示系统整体演化模式和关键行为方面发挥着至关重要的作用。随着研究的深入，人们日益认识到，现实世界复杂系统中的传播过程受到多种相互交织因素的影响，例如信息表达模式、文化背景、生活环境、认知偏好和公共政策 [19； 20]。其中许多因素难以通过固定参数或预定义规则明确纳入模型，因此，为了保持分析可处理性和计算效率，往往不得不将其舍弃 [19； 21]。对此，研究人员持续探索替代性的建模与分析方法。

随着人工智能（AI）的快速发展，尤其是近年来以 ChatGPT 为代表的大语言模型（LLM）得到广泛应用并日益受到关注，随后又出现了 Google Gemini 和 DeepSeek 等产品 [22; 23; 24; 25；26], , 人们得以探索感知和模拟人类认知、不断变化的政策措施、日常行为习惯，乃至谣言诱发的恐慌等集体反应对传播过程的影响。这为理解传播现象背后的机制与动力学模式提供了新的视角。

这些发展可归因于两个方面。首先，大语言模型以自监督方式 [27] 在多样化的文本语料库上进行大规模上下文预测训练，其文本来源涵盖广泛 [28], , 从反映人类知识积累的经典著作和权威著作，到各领域专业活动与机构活动的正式记录，以及大量捕捉日常交流、个人表达和社会交互的非正式多语言在线文本。通过这一过程，高维神经表征捕捉丰富的统计规律，并隐式整合与语言使用、现实世界语境以及人类行为模式相关的广泛跨领域知识，从而支撑自然语言理解与生成方面的高级能力。因此，大语言模型已被许多领域采用，甚至开始影响劳动力需求 [29； 30]。例如，在自然科学中，大语言模型已被用于文献综述综合、实验过程辅助和跨学科概念关联 [31]；在生物医学中，大语言模型可用于疾病知识解读、药物靶点描述和治疗建议生成 [32]；在经济学和金融领域，大语言模型可以整合来自市场叙事、政策文件和企业数据的语义信息，以实现更丰富的语境理解和战略比较 [33]；在社会环境中，它们可以参与内容生成、主题拓展和上下文感知对话 [34； 26]。在传播动力学研究中，这些能力使大语言模型成为分析与传播相关的信息和行为的多功能且有效的工具。

其次，除了作为分析工具之外，大语言模型正日益影响多个学科的研究，包括行为科学 [35], 社会学 [35; 36； 37], 政治学 [38； 39], 以及认知科学 [40], , 并开始影响当代社会的经济活动、社会实践和文化生产。从社交平台上由大语言模型控制、具有人类特征的账号，到嵌入自动驾驶车辆和数字服务中的对话代理，大语言模型驱动的智能体正迅速拓展其在传播渠道、社会互动、经济活动和日常生活中的存在。因此，复杂的社会系统已不能再被视为仅由人类智能体构成，而是日益包含彼此互动并共同演化的人类与人工智能驱动的代理所组成的异质性群体 [41]。此外，人类与大语言模型之间的互动可能形成持续的反馈回路：人类生成的数据会影响模型训练，而模型生成的内容反过来又会塑造人类后续的偏好和行为 [42]。这类人类与人工智能协同演化过程会在传播动力学中引入新型耦合与非线性，可能对信息扩散和疾病相关行为均产生影响。因此，大语言模型不仅是外部观察与分析的工具，也可能成为传播过程中的积极参与者，直接与传播动力学本身相耦合并对其施加影响 [43； 44]。

在此背景下，关于传播动力学的研究呈现出新的转变。一方面，传播的内容不再被视为抽象标签或简单载荷。相反，可以通过大语言模型所捕捉的语义、叙事模式和表达策略对其进行建模，使内容本身成为传播过程中的一个可建模组成部分 [45； 40]。例如，在基于大语言模型的基于智能体的建模中，智能体能够通过语言表达意图，并根据对多模态环境的感知生成依赖情境的行为。这使得传播过程中的行为更新不再依赖固定的转移规则，而是由认知、经验和社会情境决定，从而体现出更接近现实世界互动情境的异质性与动态不确定性 [40；19； 46]。

另一方面，大语言模型与人类或多智能体系统之间的互动会影响个体获取信息、形成理解和作出行为反应的方式，从而在一定程度上影响信息传播的路径、速度和反馈结构 [47；43；44； 48]。例如，在社交媒体上的数字信息传播中，由大语言模型驱动的智能体人群无需依赖预先设定的行为规则或大量经验校准，便能通过基于语言的推理和互动呈现丰富且依赖情境的异质性。当这类智能体日益直接地参与传播过程时，与传统的简单传染模型相比，它们会带来新的动态生成异质性形式 [49；50；51； 52] 以及复杂传染模型 [53； 54], , 这为传播动力学研究带来了新的问题与挑战。在生物流行病传播情境中，个体在面对不确定性时，可能会通过与大语言模型互动获得解释、建议或治疗方案对比，而这些反馈可能影响其行为选择。例如，在健康相关场景中，用户可以用自然语言描述症状，并获得有关风险和防护策略的建议。这可能影响他们对出行、社会互动、就医或隔离的决策，从而间接重塑人群的接触网络。

因此，有必要系统地综述并总结大语言模型与传播动力学之间的关系，以明确它们在这些过程中的作用和潜在影响。具体而言，我们撰写这篇综述有三个原因。第一，大语言模型不仅有助于揭示和分析传播机制与现象，还可能成为复杂系统传播过程中的新参与者，从而带来新型交互和扩散模式。然而，大语言模型如何有效增进我们对传播动力学的理解，以及当它们自身成为这些过程中的活跃智能体时会带来哪些新变化，目前仍不清楚；这一问题尚未得到系统考察。第二，近年来，越来越多的研究人员投入大量精力研究这一新兴领域，揭示了引人关注的现象，并取得了诸多新发现。尽管取得了这些进展，对相关研究的系统合成仍然缺乏，使新晋研究人员或感兴趣的学者难以迅速

掌握当前状况和最新进展。此外，大语言模型如何将传统的传播相关任务转变为能够发挥这些模型优势的生成式范式，从而提供更有效的分析或决策支持服务，仍是一个悬而未决的问题。同时，将大语言模型应用于传播相关研究也带来了若干尚未得到全面分析的局限、风险和挑战。

基于上述考量，本综述聚焦于两类具有代表性的复杂系统：网络空间复杂系统，涵盖谣言、错误信息和意识形态扩散等数字流行病；以及物理复杂系统，涉及流感 [55], 肺炎 [56], 登革热 [57和埃博拉 [59 60], 等生物疫情。通过考察这两个典型领域，本综述旨在推动有关大语言模型如何助力研究传播动力学的研究，并加深对大语言模型自身成为活跃参与者时传播过程内在特征的理解。本文系统梳理并归纳了近期相关研究的三个主要方面：流行病建模、流行病检测与监测，以及流行病预测与管理。图1展示了本文的架构，表1总结了利用大语言模型增强传播过程不同阶段的代表性研究。对于每项研究，表中均指出其关注的主要阶段，以及大语言模型在相应框架中发挥的功能作用。第2节介绍传统传播模型和大语言模型的背景。第3节综述基于大语言模型的流行病建模方法及其对传播过程的影响。第4节探讨大语言模型在流行病检测与监测中的应用，展示其如何感知传播对象和过程。第5节讨论大语言模型在流行病预测及支持管理方面的应用。最后，第6节指出尚待解决的问题，并概述未来研究的潜在方向。

表1：关于传播过程的LLM相关研究概述。

Type	研究	LLM	大语言模型的作用
建模	Williams等人 [61]	gpt-3.5-turbo-0301	模拟流行病传播中的人类行为
	Li 等人 [62]	gpt-3.5-turbo-1106	虚假信息传播的模型动态
	FPS [63]	gpt-3.5-turbo-1106	模拟对假新闻的态度动态
	GAG [64]	Llama-3, GPT-3.5-turbo	模型图演化
	SSF [65]	gpt-4o-mini	模拟极化和回音室效应
	Chuang等人 [66]	gpt-3.5-turbo-16k	模拟观点动力学
	FUSE [67]	gpt-4o-mini	模拟真实新闻逐渐扭曲成虚假新闻的过程 news
	SMIST [68]	gpt-4	模拟虚假信息易感性
	Mou 等人 [69]	gpt-3.5-turbo	模拟社会运动中的核心用户
	FDE-LLM [70]	ChatGLM-4	模拟意见领袖的态度与行为
	RumorSphere [71]	gpt-4o-mini	谣言传播中的核心智能体模型
	Cheng 等人 [72]	Qwen2.5-3B-Instruct, GPT-3.5	模拟公共事件动态
	EIN [73]	Gemma-2-9B	为用户回应生成立场标签
	BotSim [74]	gpt-4o-mini	构建生成并传播有害信息的用户节点
	Radivojevic 等人 [44]	gpt-4, Llama-2, Claude-2	模拟与真实在线用户的互动
	De 等人 [75]	gpt-3.5-turbo	模拟社交网络增长和拓扑形成
	Chang 等人 [76]	gpt-3.5-turbo，GPT-4o	生成社交网络
	Papachristou 等人 [77]	GPT-3.5、gpt-4	生成友谊、通信和就业网络
	Lu 等人 [78]	GPT-4o	生成对抗性社会影响内容
	Bandara [79]	-	扩大阴谋论和假新闻的影响
	Villaplana 等人 [80]	GPT-3.5	模拟持续发生的大流行病
	EpiLLM [81]	GPT-2, DeepSeek-R1, Gemma-3	对传染病传播的时间演化进行建模
	MIFlu [82]	GPT-2	融合来自不同模态的数据
	PandemicLLM [83]	Llama-2	融合来自不同模态的数据
	LLM-DG [84]	gpt	丰富患者、疾病和出院总结的表示
	Kwok 等人 [85]	gpt	支持疾病的开发和验证传播模型
	Zaslavsky 等人[86]	GPT-3.5，Claude-2	增强空间细分模拟
	Wang等人 [87]	GPT-3.5	生成个体活动轨迹数据

表1：关于传播过程的LLM相关研究概述。

Type	研究	LLM	大语言模型的作用
	Wu 等人 [88]	-	解释社会模式、构建理论，并为假设生成提供可解释的基础
	Nudo 等人 [89]	Gemini, Mistral, DeepSeek	模拟社交媒体上的政治话语
	AgentSociety [90]	DeepSeek-V3	模拟智能体之间的大规模交互，以及智能体与环境之间的交互
	Light Society [91]	gemini-2.0-flash-001, gpt-4.1-nano	建模社会信任与信息扩散
	TinyTroupe [92]	gpt-4o-mini	模拟真实的人类行为
检测	ARG [93]	gpt-3.5-turbo	提供多角度、有指导性的理由
	GenFEND [94]	GLM-4、gpt-3.5-turbo	模拟社交媒体用户并生成多样化评论
	LOCAL [95]	Flan-T5，gpt-3.5-turbo	事实核查
	SynthX [96]	gpt-4	综合虚假信息检测和解释-tion 输出
	MiLk-FD [97]	Llama-2	从新闻 con- 中获得优质的初始特征向量tent
	SR^3 [98]	Llama-3	为新闻生成正向和负向推理
	F3 [99]	gpt-3.5-turbo, Llama-2、Dolly-2、Palm-2	生成并检测虚假信息
	LESS4FD [100]	GPT-3.5、Llama-2	提取并嵌入新闻主题和真实实体
	MIL [101]	gpt	为帖子生成立场解释
	Krykoniuk 等人 [102]	T0, all-Mini-LM-L6-v2	生成预期回复
	MAGE-fend [103]	Qwen-VL，GLM-4	提供用于分类的推断理由文本
	DEEP [104]	LLaVA-13B	生成支持性和反驳性证据
	FCRV [105]	-	识别并提取主张和实体
	Zeng等人 [106]	BERT-large, DeBERTa-large, RoBERTa-large	检测错误信息
	TED [107]	gpt-4o-mini，BERT	模拟辩论过程
	FactAgent [108]	gpt-3.5-turbo	模拟人类专家核实新闻主张的行为主张
	OKI [109]	GPT、LLaVA-13B	整合开放域知识
	Costabile 等人 [110]	Llama-3.1-8B, Gemma-2-9B, Mistral-7B	检索证据、评估主张，并判定其真实性——ity
	DELL [111]	Mistral-7B, Llama2-70B，GPT	生成新闻反应，生成解释，用于代理任务，合并特定任务专家
	NIMGA [112]	gpt-3.5-turbo	完善多元视角并探究语义演变
	SePro [113]	GPT-3.5-Turbo-0125	检测谣言并生成解释
	GLPN-LLM [114]	GPT-4o	生成伪标签
监测	Deiner等人 [115]	gpt-3.5-turbo-0613, gpt-4-0314, GPT-4o-2024-05-13，Claude-3, Mixtral-8x22b，Llama-3-70B	对合成数据中的流行病学特征进行分类以及真实世界社交媒体帖子
	Deiner等人 [116]	GPT-3.5、gpt-4	根据以下内容评估流行病发生的可能性：推文
	Kaur等人 [117]	-	处理多源信息，适用于多种语言环境
	Xie等人 [118]	BERT, RoBERTa, GPT-2，BLOOM，Llama-2	识别并分析与 COVID-19 感染相关的帖子tion
	Consoli等人 [119]	gpt	提取有价值的疫情信息
	Sufi等人 [120]	gpt-4	分析新冠肺炎主题

表1：关于传播过程的LLM相关研究概览。

Type	研究	LLM	大语言模型的作用
预测	Wang等人 [121]	gpt-4-turbo	理解并筛选新闻内容，将文本信息整合到时间序列预测中
	CARE [122]	-	推断社交网络中用户的下一次激活
	Diffusion-LLM [123]	-	预测社交网络中的扩散
	Chen 等人 [124]	Qwen	预测社交媒体热度
	AutoCas [125]	GPT-2, Llama-3.2, 通义千问2.5	预测级联热度
	ComPaSC^3 [126]	Llama-3.1-8B	将视频和社区特征转换为语义增强的表示
	ALL-RP [127]	gpt-Neo	学习用户时序特征以进行转发预测
	S^3 [128]	GPT-3.5、GLM-6B	模拟社交网络中的真实人类行为
	Chuang等人 [66]	gpt-3.5-turbo-16k	模拟观点动力学
	Li 等人 [62]	gpt-3.5-turbo-1106	虚假信息传播的模型动力学
	SMIST [68]	gpt-4	模拟虚假信息易感性
	FDE-LLM [70]	ChatGLM-4	模拟意见领袖的态度和行为
	Y Social [129]	Llama-3-7B	模拟用户互动、内容传播和网络动态
	LLM-AIDSim [130]	Llama 3	模拟社交网络中的影响力扩散
	MOSAIC [131]	GPT-4o	预测用户行为，例如点赞、分享和举报内容举报内容
	MIFlu [82]	GPT-2	融合不同模态的数据
	PandemicLLM [83]	Llama-2	融合不同模态的数据
	LLM4Cast [132]	TinyLlama	预测病毒性疾病趋势
	LLM-DG [84]	gpt	丰富患者、疾病和出院总结的表示
	Priya等人[133]	Medisearch AI	执行主题建模、情感分析等功能，以及关系抽取
管理	Almaliki 等人[134]	gemini-1.5-pro-latest	处理用户评论，识别网络仇恨实例，并生成适当的响应
	Jo等人[135]	HyperCLOVA	提供自然语言理解和自然语言生成能力

2. 背景

本节首先综述传播动力学的经典模型，包括流行病学模型和社会传染模型，以介绍基础背景。随后，系统概述大语言模型的核心技术与理论基础，分析主流模型架构的关键特征，并重点介绍大语言模型在表示与传播过程相关的文本与上下文特征方面所具备的独特能力。表2展示了本文中使用的符号和记法，表3则概述了大语言模型的关键基础概念。

2.1. 传播动力学入门

2.1.1. 基本概念

为便于理解传播动力学，我们简要介绍网络科学和传播动力学中的几个基本概念。网络为交互实体提供结构化表示，说明谁能够影响谁。在这一结构基础上，传播动力学指状态、信息或行为通过互动随时间传播的过程。网络定义了互动路径，而传播动力学则描述状态沿这些路径的时间演化，从而形成大规模扩散或消亡等系统层面的行为。

- 网络：网络为由交互实体构成的系统提供了一种抽象的数学表示。形式上，网络定义为图 $G = (V,E)$，其中 V 表示有限的节点（或顶点）集合，而 $E \subseteq V \times V$ 表示由边组成的集合，用于编码节点之间的成对关系。网络的结构连通性也可以通过邻接矩阵 $\mathbf{A} = [a_{ij}]$，来等效描述，其中每个元素 a_{ij} 表示节点 i和j之间是否存在边；如果两个节点相连，则$a_{ij} = 1$，否则$a_{ij} = 0$ [136；18； 137]。

表 2：符号和记法。

符号	描述
$S(i)$，$I(i)$，$R(i)$	个体i的易感、感染和康复状态
ρ_S, ρ_I, ρ_R	时间 t 时易感者、感染者和康复者所占的节点比例
β	传播（感染）率
γ	康复率
N	总人口规模
w_i	序列中的第i个词元
$x = \{w_1, \ldots, w_{\|x\|}\}$	输入词元序列（提示或观测文本）
$y = \{w_{\|x\|+1}, \ldots, w_{\|x\|+T}\}$	生成词元序列（响应）
$\|x\|$	输入序列长度 x
T	生成的响应序列长度 y
$\mathcal{V}$	所有可能词元的词汇表
d	嵌入或隐藏维度
$\mathcal{D}$	训练数据集或语料库
θ	模型参数
$P_\theta(x)$	模型参数 θ 下序列 x 的概率分布
$P(w_i \mid w_{<i}; \theta)$	在给定先前词元$w_{<i}$的条件下，词元w_i 的条件概率
$\mathcal{M}$	将提示词映射为响应的大语言模型
$\mathbf{h}_t \in \mathbb{R}^d$	位置或层t处的隐藏状态向量
$\mathcal{L}$	大语言模型训练中使用的损失函数（例如交叉熵）

• 网络的关键属性：

– 度：节点的度表示其所拥有的连接数量。– 平均度：它是网络中所有节点度的平均值，用于粗略衡量网络的整体连通性水平。– 度分布：它描述网络中不同度值出现的频率，是衡量网络连接异质性的重要统计特征。– 网络密度：它是网络中实际连接数与最大可能连接数之比，用于衡量网络中节点之间关系的紧密程度，反映整体网络连接的稠密程度。– 平均路径长度：它是所有节点对之间最短路径长度的均值，反映全局连通的效率。– 枢纽节点：在网络中，枢纽节点是连通性异常高的节点，能够对传播过程产生不成比例的影响力。– 社团/簇：它指网络中紧密连接的节点子集，其中传播过程通常在社区内部迅速传播，然后通过有限的跨社区连接向外扩散。

• 网络拓扑：网络拓扑刻画节点之间连接的全局组织方式，并为传播过程提供其赖以发生的结构基础。在复杂系统中，传播通常在网络$G = (V, E)$上展开，其拓扑决定哪些接触是可行的，以及影响力如何在系统内传递。连通性、异质性和时间特征的变化会显著改变宏观传播行为，包括流行病阈值、传播速度等。经典网络拓扑包括随机网络、无标度网络、小世界网络、规则网络等。下面，我们依次介绍这些具有代表性的网络结构。

– 随机网络：指在给定固定节点集的情况下，以预设概率在任意节点对之间独立生成边的网络。由于边的形成机制具有随机性，节点的度通常集中在平均值附近，因此结构异质性相对较弱。这一特性通常与有限且非零的传播阈值相关。 [138].– 无标度网络：度分布遵循幂律分布的网络。在无标度网络中，节点的连通性（度）具有高度异质性：少数节点的度极高，而大多数节点只有少量连接 [139]。无标度网络的度分布具有高度异质性，通常会导致阈值趋于零，并增强传播潜力 [138]。– 小世界网络：一种兼具较高局部聚类和较短平均路径长度的网络，使信息或疾病能够在局部快速传播，同时也能通过少量长程连接，高效传播至网络的远端区域。

– 规则网络：由大量相同或相似的节点组成，这些节点通过规则或固定的连接相互交互。在规则网络中，每个节点都遵循相同的计算规则，并与其邻近节点交互 [140]。– 多层网络：由一组共享节点构成，但节点之间通过不同类型的连接关系相连，从而形成不同的网络。 [141].– 时序网络：一类明确考虑连接随时间变化的网络结构，其中边仅在特定时间段内有效 [142]。通过纳入时间信息，时序网络能够更真实地反映接触顺序和时间约束对传播路径及传播效率的影响。– 高阶网络：复杂网络的一种替代表征方式，能够明确建模超越成对边的交互，并日益广泛地应用于复杂系统研究 [143；144；145；146；147； 148]。

• 节点状态：节点状态表示参与传播过程的实体所处的离散状态，例如易感、感染者、康复者或已采纳。

• 状态转变：描述节点状态如何通过与邻居的交互或内部动力学发生变化，例如感染、采纳、康复或遗忘。

• 传播率：传播率 λ 用于表征传染过程的有效强度，通常以传播率 β与康复率 γ [149]之比来量化。

• 流行病阈值：流行病阈值表示疾病或信息能够在网络中广泛传播的临界条件。一般而言，只有当病原体的传播率超过流行病阈值 λ_c 时，才会传播。随机网络具有有限的流行病阈值 λ_c，这意味着传播率较低的病原体（$\lambda < \lambda_c$）必然会消亡。然而，如果病原体的传播率超过 λ_c，病原体就会形成地方性流行，并且在任何时刻都有一定比例的人群受到感染。对于无标度网络，我们有 $\lambda_c = 0$，因此，即使传播率很低的病毒 λ 也能在人群中持续存在。

• 传播速度：指疾病、信息或行为在网络中传播的速度。

• 级联：级联是由传染过程产生的一系列激活，其中节点以一定概率使相连的节点被激活。类比于传染病在网络中的传播，一个感染（已激活）的节点会使其粉丝暴露于感染之中。随着暴露的粉丝被感染，疾病便在网络中级联传播，这些粉丝继而使自己的粉丝暴露于疾病之中，如此不断延续 [150]。

2.1.2. 传播动力学经典模型

根据传播实体的性质，复杂系统中的传播动力学模型可分为两大类：用于**生物流行病**的**流行病学模型** ，用于描述病原体在生物个体之间的传播；以及用于数字流行病的**社会传染模型** ，侧重于数字信息在社交网络中的扩散，例如观点、谣言和计算机蠕虫。前者源于公共卫生和传染病研究，强调生理感染—恢复过程，通常建立在接触网络上的质量作用原理或马尔可夫假设之上。后者则源于社会学和计算机科学，着重关注“感知—决策—交互”的社会认知机制，通常基于阈值规则或概率级联过程。这两类模型在建模假设、动力学性质和传播机制方面存在根本差异。下文将分别介绍。

流行病学模型。 对于生物疫情，传播主要通过身体接触发生，感染概率通常与接触频率相关。这类模型通常基于均匀混合假设，即个体之间的接触是均匀的，或可通过接触网络建模 [149; 151]。个体通常用有限的一组状态表示，例如易感者（S）、感染者（I）、康复/移除者（R）和暴露者（E）状态。这些状态之间的转变遵循特定的疾病机制。其动态通常使用常微分方程或连续时间马尔可夫过程建模。通常，不同类型的生物疫情需要不同的模型，其中最具代表性的模型是 SIS 和 SIR 模型：

- **SIS模型**：该模型描述不具有永久免疫力的疾病（例如普通感冒）。个体要么是易感者，要么是感染者。其动态反应过程为：

$$S(i) + I(j) \xrightarrow{\beta} I(i) + I(j), \quad I(i) \xrightarrow{\gamma} S(i), \tag{1}$$

其中，i 和j 表示不同的个体，β 为有效接触率，γ 为康复率。易感者$S(i)$与感染者$I(j)$接触后会被感染，而感染者以γ的速率康复并返回

易感状态，形成$S \rightarrow I \rightarrow S$循环。在质量作用假设下，相应的动力学演化方程如下：

$$\begin{aligned} \frac{d\rho_S}{dt} &= -\beta\frac{\rho_S\rho_I}{N} + \gamma\rho_I, \\ \frac{d\rho_I}{dt} &= \beta\frac{\rho_S\rho_I}{N} - \gamma\rho_I. \end{aligned} \tag{2}$$

此处，$N = \rho_S + \rho_I$表示总人口。ρ_I和ρ_S分别表示时间t时易感个体和感染个体的比例，反映了群体层面的感染动力学。术语$\beta\rho_S\rho_I/N$量化了均匀混合条件下单位时间内预期的新感染率。

- **SIR模型**：适用于具有永久免疫力的疾病（例如天花）。SIR模型中的个体处于三种S、I和R状态，动力学反应过程为：

$$S(i) + I(j) \xrightarrow{\beta} I(i) + I(j), \quad I(i) \xrightarrow{\gamma} R(i), \tag{3}$$

这意味着易感个体以β的速率变为感染者，而感染个体以γ的速率康复并进入康复状态R，形成单向的$S \rightarrow I \rightarrow R$路径。当$N = \rho_S + \rho_I + \rho_R$时，动力学演化遵循：

$$\begin{aligned} \frac{d\rho_S}{dt} &= -\beta\frac{\rho_S\rho_I}{N}, \\ \frac{d\rho_I}{dt} &= \beta\frac{\rho_S\rho_I}{N} - \gamma\rho_I, \\ \frac{d\rho_R}{dt} &= \gamma\rho_I, \end{aligned} \tag{4}$$

其中，ρ_R表示时间t时康复个体的数量。

社会传染模型。尽管社会传染和流行病传播都是在网络上展开的扩散过程，但它们在个体层面依赖不同的行为假设。经典流行病学模型通常将传播具体化为接触所导致的暴露函数：原则上，单次传染性接触就足以引发感染，而额外的接触主要会提高传播概率。相比之下，许多受社会影响的采纳行为并非由单次暴露触发。简单传染（如简单信息和熟悉的观点）可以通过单次接触传播。而复杂传染（如职场文化的变化、在线社交平台的使用以及创新技术的采纳）通常往往需要社会强化，即在改变观点或采纳行为之前，反复暴露于相关信息和/或获得多位同伴的佐证；当相关行为代价高、风险大或陌生时尤其如此 [10；11； 152]。

这类传播过程强调决策机制，通常采用概率规则或博弈论规则，可分为阈值模型和独立级联模型 [153； 154]。基于阈值的模型通常与复杂传染相关，因为其中的采纳被建模为一种集体响应现象：只有当采纳行为的邻居数量或比例超过某个阈值时，个体才会变为活跃节点，从而明确体现多次接触带来的强化。相比之下，独立级联（IC）模型通常被视为简单传染过程，在其中，每次暴露都会独立发挥作用，并可能以一定概率触发采纳，因此其动态过程类似于生物病毒的传播。

- **阈值模型**：如果已激活邻居的比例超过个人阈值，个体就会采纳或传播信息。形式化地，设$a_v \in [0, 1]$为节点v的激活阈值，f_v为其邻居中活跃节点所占的比例。那么，当$f_v \geq \theta_v$时，v就会变为活跃节点。此过程通常是确定性的。

- **独立级联模型**：一种概率模型，其中每个活跃节点u都会尝试以激活概率$f(u, v)$激活每个不活跃邻居v，且该概率独立于先前的尝试及其他邻居。每一轮中，活跃节点都会尝试激活邻居，无论结果如何，之后的轮次中都不会再次尝试。该过程会持续进行，直至不再出现新的激活。

2.2. 大语言模型简介

大语言模型是一类先进的人工智能系统，通过海量自然语言语料库进行训练，以捕捉潜在的统计模式、句法结构和语义关系，从而生成与上下文相关且连贯的自然语言输出 [155；156； 157]。大语言模型通常基于Transformer架构构建，利用大规模训练数据实现出色的跨任务迁移和泛化能力 [23； 155]。这些能力不仅使大语言模型能够执行文本生成、对话系统和知识检索等传统任务，还能直接或间接地支持与复杂系统中流行病传播动力学相关的任务。例如，大语言模型可以直接替代基于主体的模型中的基于规则的智能体，从而更逼真地模拟个体交互和传播路径。同时，它们还可以通过增强对多模态流行病相关数据（如图像、文本报告和时间序列）的感知和表示，间接支持流行病建模和预测，进而提高准确性和鲁棒性 [19；40；158；83；159；160；161； 162]。通常，核心学习过程的

表 3：有关大语言模型的重要概念

Transformer 是一种神经网络架构，利用自注意力对序列数据中的全局依赖关系进行建模。它并行处理序列中的所有元素，使每个词元都能动态关注其他词元。这种设计有助于高效学习长距离上下文关系，并已成为大多数大语言模型的基础。
词元是模型处理的最小意义单位，可以是单词、子词或字符，具体取决于模型分词过程的粒度。
嵌入是将词元映射到连续空间的向量表示，从而以数值形式捕捉语义和句法信息。
提示词是提供给大语言模型的输入或指令，用于引导其输出生成，并决定其在特定任务中的行为。
提示工程是指设计和优化提示词，以引导大语言模型生成期望的响应，通常无需修改模型参数。
零样本学习 / 少样本学习是指大语言模型在没有（零样本）或仅有少量（少样本）特定任务示例的情况下，泛化到新任务的能力。 **微调**是指使用新的数据集或任务对预训练模型或已训练模型进行再训练。 **多模态模型** 是指能够理解并整合来自多种 模态（如文本、视觉或听觉数据）的模型。

大语言模型的目标是估计给定输入序列的概率分布。形式上，其训练优化目标可以表示为最大化相对于模型参数的序列对数似然，如下所示：

$$\max_{\theta} \sum_{i=1}^{|x|} \log P(w_i \mid w_{<i}; \theta), \tag{5}$$

其中，w_i 表示输入序列中的第i个词元$x = \{w_1, \ldots\ldots, w_{|x|}\}$ 而 θ 指模型参数。通过调整参数，大语言模型旨在最大化生成正确序列的概率。

2.2.1. 使大语言模型具备功能的过程

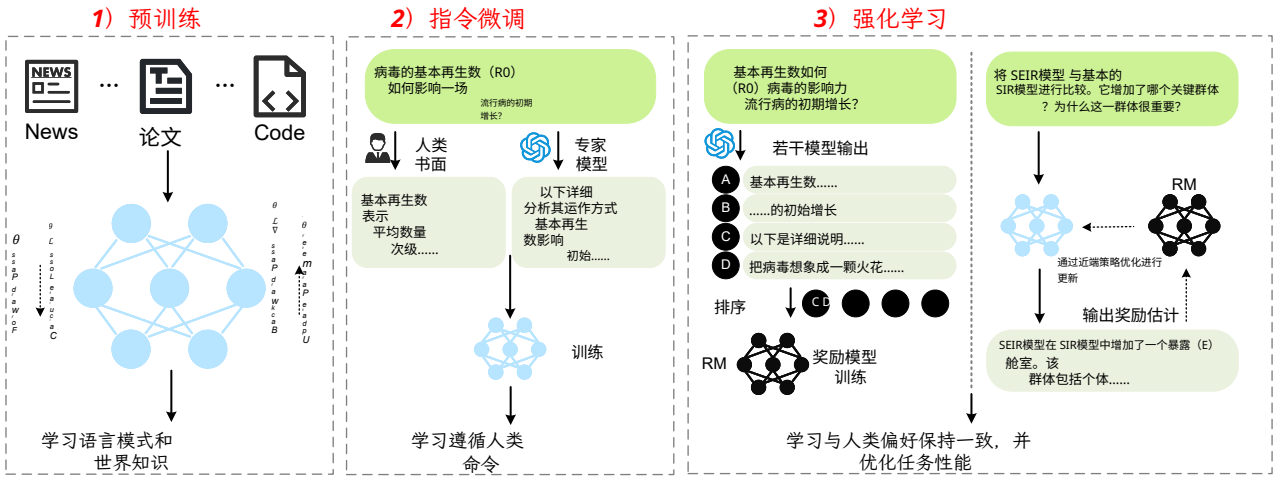


图2：如何使大语言模型能够处理各种复杂任务的示意图。(1) 预训练：使用海量数据集进行自监督学习。其目的是教会模型掌握语言规则和知识，并形成基础模型。(2) 指令微调：使用指令数据进行监督训练。其目的是教会模型在对话过程中理解并遵循人类指令。(3) 强化学习：使用偏好数据进行进一步优化，以确保输出有用、安全并符合预期。

为了使模型逐步学会完成各种复杂任务所需的最优参数，大语言模型通常要经历三个阶段：预训练、指令微调，以及基于人类或人工智能反馈的强化学习，如图 2 所示。

在预训练阶段，模型通过在大规模文本数据（如新闻文章、学术论文和代码）上进行自监督学习，掌握语言的基本结构和世界知识 [163； 156]。这些数据被输入神经网络，模型通过前向传播计算预测误差，并用损失函数 $\mathcal{L}(\theta)$ 表示。随后，模型的参数 θ 通过反向传播进行更新，从而得到预训练大语言模型。

在指令微调阶段，预训练模型将进一步学习如何根据给定指令执行任务。该阶段使用由人类书面和专家模型生成的问答

对组成的训练数据集 [164; 165; 166]。例如，模型将在此阶段学习回答诸如“病毒的基本再生数如何影响流行病的早期增长？”之类的具体问题。这种微调使模型能够根据当前任务调整其输出，从而提高其在复杂系统中准确响应各种需求的能力。

在强化学习阶段，大语言模型的性能通过反馈信号 [167 得到优化； 168]。首先，人工评估者会根据质量对模型生成的多个候选回答进行排序。这些排序信息用于训练奖励模型，使其根据人类偏好为不同回答分配奖励 [169]。在这一阶段，奖励模型会对每个生成的响应进行评分，并向大语言模型发送反馈，使其能够相应地调整输出。例如，在生成流行病管理策略的任务中，模型首先会生成几条管理建议，每条建议对流行病传播的潜在影响各不相同 [170； 171]。接下来，人工评估者会根据这些策略的科学有效性、准确性以及与实际数据的一致性进行排序。利用这些反馈，训练奖励模型为每条新建议分配奖励。在这一过程中，大语言模型会根据奖励模型的反馈不断调整策略生成方式，从而优化其生成的管理策略。最终，通过这种强化学习机制，模型能够更好地应对复杂的流行病传播问题，并提供符合现实需求的管理策略。

2.2.2. 大语言模型在流行病相关任务

完成三个训练阶段后，大语言模型便具备处理各种复杂任务的能力。通过调整输入形式和输出格式，可以灵活地将其应用于流行病相关场景。例如，在流行病建模中，模型可以根据不同的提示词生成具有特定角色或功能的智能体；在监测任务中，则可以按照预定义格式生成结构化的早期预警信号。通过这种输入与输出的定制，大语言模型能够将其在语言理解和知识建模方面的通用能力转化为研究传播动力学的具体功能。

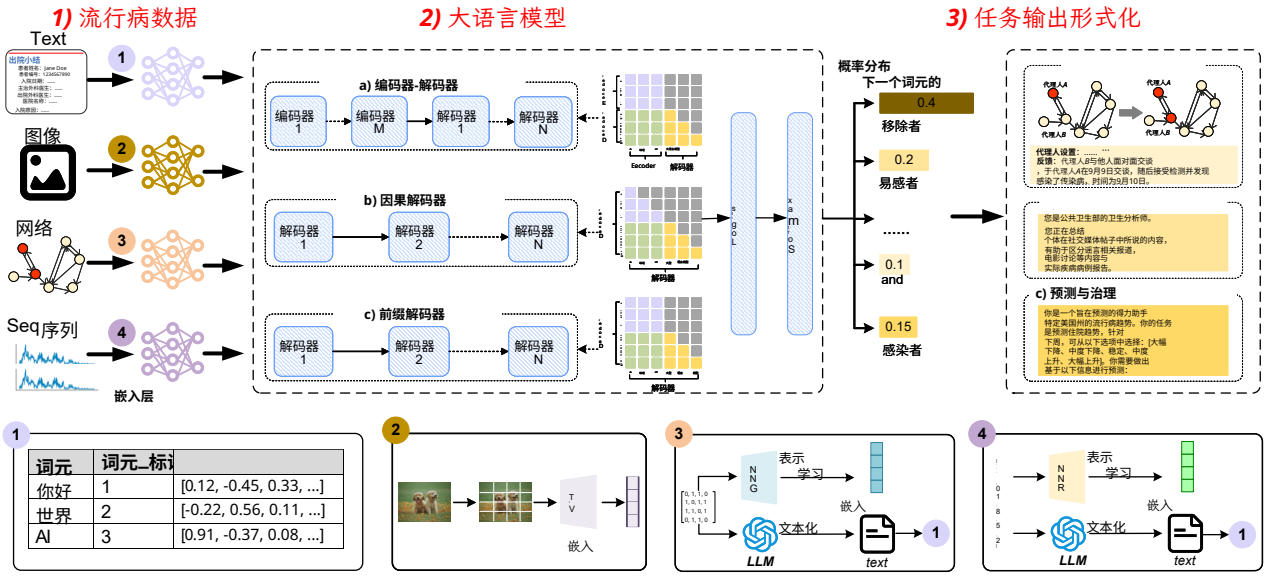


图 3：大语言模型工作流程示意图。(1) 数据嵌入：各种类型的输入数据，包括文本、图像、网络和时间序列，都会转换为称为嵌入向量的高维向量，以便大语言模型进行处理。(2) 大语言模型：大语言模型利用采用编码器-解码器、因果解码器或前缀解码器等 Transformer架构的神经网络处理嵌入向量，以预测下一个词元。(3) 任务输出形式化：最后一步是针对不同任务定制模型的输出格式，确保大语言模型能够适用于各种应用场景，例如流行病建模、监测、检测、预测和管理。

在大语言模型处理文本、图像、图结构和序列数据等多模态输入时，原始输入首先会被转换为高维向量表示，即嵌入。这些向量可以视为坐标，用于捕捉输入数据的语义或结构特征。随后，嵌入经过多个Transformer层处理，生成Logits；Logits可以理解为未经归一化的分数，用来表示各个可能结果的相对偏好。接着，使用softmax函数将这些Logits归一化为词汇空间上的概率分布。最后，采用解码策略（例如top-k 采样、top-p 采样或束搜索），从这些概率分布中选取特定词元作为输出。选出的输出随后会被添加到输入中，并重复这一过程，逐步生成下一个词元，直至满足停止条件。

如图所示　图3的左侧和底部所示，不同模态的数据嵌入过程各不相同 [172; 160； 83]　　：

- **文本**：疫情相关研究通常涉及处理文本数据，例如疾病监测报告或患者症状日志。首先将输入文本切分为词元，然后将这些词元映射为词元_ID，最后通过文本嵌入层生成相应的嵌入向量。

• **图像**：在流行病数据中，图像包括用于检测疾病症状的医学影像（如CT扫描）、用于追踪疾病传播模式的卫星图像等。这些图像通常会被划分为固定大小的像素块，例如 16×16 分块，并通过视觉Transformer等视觉编码器提取特征，生成图像嵌入向量 [173] 。

• **网络数据**：在流行病建模中，网络数据可以表示疾病在社会网络中的传播，其中节点代表个体，边代表互动或传播事件。处理网络数据主要有两种方法。一种基于图神经网络（GNNs） [174], ，通过随机游走或注意力机制，从邻接矩阵和节点特征中学习低维节点或网络表示。另一种方法基于大语言模型，将网络结构转换为文本描述，再通过文本嵌入层进行处理以生成嵌入 [172]。

• **序列数据**：在流行病的背景下，序列数据可能包括随时间变化的报告病例数时间序列数据、人类移动行为，或有关疾病传播或疫苗接种率的每日报告。处理序列数据的一种方法是使用循环神经网络（RNN），通过时间层逐步编码序列模式。另一种方法基于大语言模型，将序列数据转换为文本描述，再通过文本嵌入层进行处理 [83]。

为了映射从输入嵌入到输出的转换，大语言模型通常采用基于 Transformer 设计的因果解码器、前缀解码器或编码器-解码器之一。如图3中部所示，GPT [22] 和 LLaMA [157] 系列等大语言模型采用因果解码器框架，使用单向掩码自注意力，使模型能够以自回归方式生成文本，也就是说，每个词元只能关注前面的词元。相比之下，由前缀解码器构成的大语言模型（例如 ChatGLM [175]）使用双向注意力处理输入信息，同时保持输出生成的单向解码。这种设计使模型能够有效处理上下文信息，同时避免生成过程中的冲突。其他大语言模型，例如 T5 [176]和 T0 [177], ，采用编码器-解码器架构，将编码器中的双向自注意力与解码器中的交叉注意力相结合，以捕捉输入序列与输出序列之间复杂的依赖关系。

如图3右侧所示，当大语言模型应用于复杂系统中的传播动力学任务时，通常会利用生成式输出，将各种任务的预期结果转换为文本描述。例如，在流行病建模任务中，个体节点的属性，如年龄、性别和社会关系，会以自然语言进行描述，例如“你是一名35岁的女教师，生活在一座大城市，倾向于自由派政治，重视教育平等和环境保护.....”。该文本描述被视为模型输入x，随后传递给大语言模型。基于这一输入，大语言模型会生成相应的文本响应y，反映节点状态的变化，例如“人物A与人物B接触后被感染”，从而捕捉个体之间传播不断演变的动力学过程。在推理过程中，大语言模型接收输入提示x，并以自回归方式生成响应序列 [164]。具体而言，给定初始输入提示x_i，模型会根据输入和此前生成的上下文，依次预测下一个词元，从而生成响应。生成输出序列$\mathbf{w}_{>|x|} = \{w_{|x|+1}、\ .....、\ w_{|x|+T}\}$的概率可表示为

$$P_\theta(\mathbf{w}_{>|x|} \mid x_i) = \prod_{t=1}^{T} P_\theta(w_{|x|+t} \mid w_{<|x|+t}, x_i), \tag{6}$$

其中 $w_{|x|+t}$ 表示在输入提示 x_i 和此前生成的所有词元$w_{<|x|+t}$条件下，在第 t 步生成的词元。

与将传播机制硬编码并依赖结构化变量的经典舱室模型或社会传染模型相比，大语言模型在三个方面提供了一种互补范式：（a）表示：通过提示和检索摄取异质、高维且结构较弱的信号（例如帖子、报告和临床笔记），无需专门进行特征工程；（b）适配：通过少样本或零样本泛化、工具使用和反事实角色条件化，适应非平稳性和快速的机制转换；（c）决策支持：生成便于人类理解的推理依据、政策选项和多智能体叙事，将微观行为与宏观动力学联系起来。这些差异为早期预警、监测和干预设计带来了明显优势，同时也引发了关于可靠性、偏见和管理的开放问题，而传统误差指标无法充分涵盖这些问题。在接下来的章节中，我们综述基于大语言模型的传播动力学研究现状，按照领域（数字流行病与生物流行病）和任务类别（建模、感知和监管）对相关文献进行梳理。我们将重点介绍这些研究的方法、主要发现，以及它们为流行病传播分析带来的独特优势。

2.3. 用于验证的经验数据

本节将 介绍如何使用经验数据验证基于大语言模型的传播仿真，以及主要评估 指标和所采用的数据集。

总体而言，经验数据在基于大语言模型的传播模拟中的作用可分为三个方面：初始化模拟、驱动模拟和评估模拟。首先，经验数据可用于确定传播模拟的初始条件，例如社会网络结构、用户行为特征，

以及时间序列数据。Zhang 等人 [130] 利用斯坦福网络分析项目（SNAP）小组的 ego-Facebook 数据集子集来表示网络结构。该数据集中，节点总数（代表个体用户）为 333，边总数（代表用户之间的连接或关系）为 2,500。Törnberg 等人 [178]使用美国全国选举研究（ANES）的数据创建角色画像，以指导其模型中每个智能体的行为。具体而言，他们利用这些数据确定每个智能体所接触的新闻来源、人口统计特征、政治信念和兴趣、对特定群体和政治人物的态度，以及某些非政治兴趣。

此外，经验数据不仅可以作为模拟的起点，还可以动态引导生成过程，使传播行为更加真实。Zheng等人 [179] 将时间引擎引入模拟中。该模块根据智能体的时间特征激活智能体，使其能够执行评论、发布内容以及与其他智能体和环境互动等各种行为。这些时间特征源自真实的行为数据，使整体传播过程更加符合用户在现实社交媒体平台上的日常活动模式。Yao 等人 [70] 使用 Python 开发了一个模拟社交媒体平台。他们以现实世界中真相暴露的时间为参考点，在该时刻向智能体发布包含真相的离线新闻。

最后，经验数据对于验证模拟传播过程是否忠实地再现观测到的动态同样至关重要。在此语境下，验证并不局限于点对点的准确性，而是旨在从多个层面评估模拟结果与现实世界观测结果之间的一致性，包括个体状态、时间轨迹以及总体传播趋势。因此，研究采用了多种评估指标。具体而言，准确性和宏平均 F1 分数等面向分类的指标用于评估离散状态预测的正确性；包括动态时间规整（DTW）在内的时间对齐指标用于评估模拟传播轨迹与观测传播轨迹在时间维度上的相似性；皮尔逊相关系数和余弦相似度等基于相关性和向量的度量则用于从更宏观的层面量化模拟趋势与经验趋势之间的一致性。以下给出这些指标的计算方法。

- 准确性：在多分类场景中，准确性衡量所有类别中被正确预测的样本总体占比，定义如下：

$$\text{Accuracy} = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}(\hat{y}_i = y_i) \tag{7}$$

其中，N 表示样本总数，$\hat{y}_i$ 和y_i 分别表示第i</style}个样本的预测标签和真实标签，$\mathbb{I}(\cdot)$是一个指示函数：条件为真时取值为 1，否则取值为 0。

- 宏平均 F1：在多分类场景中，宏平均 F1 通过分别计算每个类别的 F1 分数，再对各类别进行等权平均来评估模型性能，定义如下：

$$\text{Macro-F1} = \frac{1}{C}\sum_{c=1}^{C}\text{F1}_c \tag{8}$$

其中，C 表示类别总数，F1_c 是类别c的 F1 分数。

- 动态时间规整：动态时间规整用于衡量模拟态度序列S 与真实态度序列T之间的相似性，尤其适用于两个序列长度不同或存在时间变化的情况。动态时间规整通过计算最小累积距离来对齐这两个序列。公式如下：

$$DTW(S,T) = \min\sqrt{\sum_{k=1}^{K}\left(s_{i_k} - t_{j_k}\right)^2} \tag{9}$$

其中（i_k, j_k）表示序列 S 和T中的对齐点，K 是对齐路径的长度。DTW 距离越小，模拟态度与真实态度的时间序列之间的相似度越高，这表明模型在模拟复杂时间序列模式方面的表现越好。

- 皮尔逊相关系数：皮尔逊相关系数用于衡量模拟态度序列 $S = (s_1, s_2, \ldots, s_n)$ 与真实态度序列 $T = (t_1, t_2, \ldots, t_n)$ 之间的线性相关性。公式如下：

$$r = \frac{\sum_{i=1}^{n}(s_i - \bar{s})(t_i - \bar{t})}{\sqrt{\sum_{i=1}^{n}(s_i - \bar{s})^2}\cdot\sqrt{\sum_{i=1}^{n}(t_i - \bar{t})^2}} \tag{10}$$

其中，$\bar{s}$ 和 $\bar{t}$ 分别是模拟态度和真实态度的均值。皮尔逊相关系数的范围为 $[-1, 1]$，其值越接近 1，模拟态度与真实态度之间的线性相关性越强，表明模型越能准确捕捉真实态度的变化。

• 余弦相似度：余弦相似度指标反映模拟内容与真实内容之间的语义相似度，计算方法如下：

$$\text{Cos-Sim}(\mathbf{s}, \mathbf{r}) = \frac{\mathbf{s} \cdot \mathbf{r}}{\|\mathbf{s}\|_2 \, \|\mathbf{r}\|_2} \tag{11}$$

其中，$\mathbf{s} \in \mathbb{R}^d$ 和$\mathbf{r} \in \mathbb{R}^d$ 分别表示模拟内容和真实内容的d维嵌入向量。 $\|\mathbf{s}\|_2$ 和 $\|\mathbf{r}\|_2$ 分别表示对应向量的欧几里得范数。该指标反映了模拟内容在语义层面与真实内容的匹配程度。

例如，Mou 等人[69] 同时考虑微观层面和宏观层面的评估，分别关注个体用户对齐和整体系统层面结果。在微观层面，他们为每个核心用户代理提供真实的上下文信息，并在单轮模拟中评估其决策行为，重点关注立场对齐、内容对齐和行为对齐。具体而言，立场对齐通过立场分类的准确性和宏平均 F1 分数，以及态度分数的平均绝对误差（MAE）来衡量；内容对齐通过内容类型分类的准确性、宏平均 F1 分数，以及模拟内容与真实内容之间的余弦相似度来评估；行为对齐则依据发帖和转发行为分类的准确性及宏平均 F1 分数进行评估。在宏观层面，他们量化完整多轮模拟中的态度分布，同时考虑静态态度分布和平均态度的时间序列。具体而言，静态态度分布通过偏见（平均态度偏离中性态度的程度）和多样性（态度的标准差）来衡量；平均态度的时间序列则通过动态时间规整（DTW）和皮尔逊相关系数评估其与真实数据的相似度。Yao 等人[70] 也使用 DTW 和皮尔逊相关系数评估模型在模拟中的有效性。此外，Zhang 等人 [130] 使用 BERT，计算从现实世界评论中提取的主题与模拟生成的主题在多次迭代过程中的语义相似度变化。在当前研究中，使用经验数据进行验证仍然有限，但表 4 所示的候选数据集可作为潜在的验证数据来源

.

表 4：用于验证基于大语言模型的传播模拟的候选数据集。

数据集	领域	规模
Twitter15 [180]	谣言	842个事件
Twitter16 [181]	谣言	5,656个事件；4,907,641条帖子；3,238,047名用户
PHEME [182]	谣言	5个事件；5,802条帖子
RumourEval 2017 [183]	谣言	297个事件；7,100条帖子
RumourEval 2019 [184]	谣言	446个事件；7,995条帖子
MR^2 [185]	谣言	14,700个事件；1,073,487条帖子；107,892名用户
FakeNewsNet [186]	假新闻	23,196 篇新闻文章；690,732 条帖子
微博21 [187]	假新闻	9,128 篇新闻文章
MC-Fake [188]	假新闻	27,155 篇新闻文章； ~ 5M 个帖子； ~ 2M 名用户
FakeSV [189]	假新闻	738 个事件；5,538 个新闻视频
FTT [190]	假新闻	-
TruthSeeker [191]	假新闻	1,400 篇新闻文章； ~ 186,000 条帖子
MCFEND [192]	假新闻	23,789 篇新闻文章；170,713 条帖子；803,779 个用户
FauxBuster [193]	错误信息	917 个帖子；1,988,493 条评论；621,983 位用户
MuMiN [194]	错误信息	~ 26k 个讨论串； ~21M 个帖子
NETINF [52]	信息扩散	~172M 篇新闻文章和博客帖子
DeepHawkes [195]	信息扩散	55,108 条帖子
新冠-19 [196]	信息扩散	~8M 条评论和帖子
Weibo-COV [197]	信息扩散	~40M 条帖子； ~20M 个用户

3. 流行病建模中的大语言模型

传播动力学长期以来一直是统计物理学的核心主题，涵盖渗流、扩散以及网络上的输运等过程。这些框架为理解局部交互如何产生大规模集体行为提供了基础性洞察力，也为从信息扩散到流行病暴发等多种现象提供了有力的抽象模型。在这一框架中，流行病传播被建模为发生在个体相互作用网络上的动态过程，类似于状态或粒子在物理介质中的传播。传统的流行病建模方法，如舱室模型和基于主体的仿真，在很大程度上继承了物理学中确定性的、基于规则的范式，依赖固定参数和均质性的简化假设。这些模型通过设定传播率和康复率来描述流行病传播，而个体行为往往

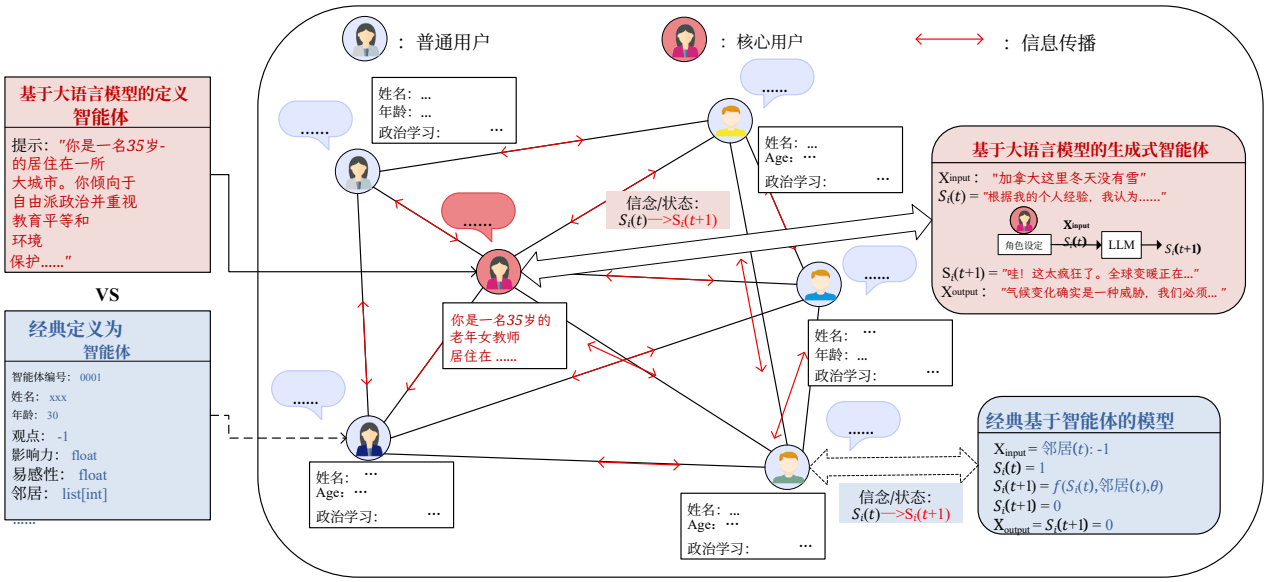


图4：经典流行病建模与基于大语言模型的流行病建模比较。在经典的基于主体的模型中，每个智能体都由一组预定义属性刻画，例如影响力、易感性和网络邻居；状态转变则根据这些参数和局部交互，通过明确规定的规则或方程计算。相比之下，基于大语言模型的生成式智能体被建模为认知驱动的个体，其状态转变源自内部推理过程和情境感知。这些智能体并非遵循固定的转变方程，而是解读环境线索并生成依赖情境的行为，从而使传播动力学和个体响应能够灵活适应认知、社会和环境因素。

被简化为平均场理论下的统一响应。尽管这类方法在解析上易于处理，却往往难以捕捉现实流行病情景中的异质性、随机性和多模态影响 [19; 61]。

相比之下，大语言模型具备类人内部表征能力，能够自发形成与人类认知高度相似的概念系统，从而有助于捕捉个体复杂的类人内部特征 [198；19；199； 200]。如图4所示，在基于大语言模型的流行病模型中，每个智能体的状态转变并不依赖明确的方程或参数化规则，而是取决于智能体类人的内部认知过程及其对外部环境变化的感知。同时，凭借强大的理解和推理能力，基于大语言模型的智能体能够动态感知多模态外部环境，包括文本、图像和网络数据，并通过自然语言推理生成依赖情境的行为响应。在这一框架下，疾病传播率和个体行为模式会根据认知过程、过往行为、文化背景等动态调整 [61]。这种方法将个体行为预测从固定规则转向由多重因素相互作用所塑造的实时生成，从而更好地反映现实世界流行病的复杂性和不确定性。通过使智能体能够在社会、文化和环境线索的影响下做出适应性决策，基于大语言模型的建模能够更真实地再现复杂人群中的流行病传播动力学。表5从多个方面总结了流行病的基于大语言模型的智能体建模方法，包括智能体相关设置（大语言模型、规模和记忆）、网络结构、状态表示和交互机制。在接下来的小节中，我们将分别综述大语言模型在数字流行病和生物流行病建模中的应用。

3.1. 基于大语言模型的数字流行病建模

在谣言、错误信息、阴谋论或观点等数字信息传播的流行病学建模中，感染者是指已接收并持有某条特定信息，且能够将其传递给他人的智能体。易感者是指尚未接触到该信息，但可能在暴露后受到感染的个体。康复者被视为对该信息免疫，这意味着他们即使接触到该信息，也不会受到其影响。例如，在谣言传播中，感染者是指听说并相信谣言的人，易感者尚未接触到谣言，而康复者则已认清谣言的虚假性。

近年来，大语言模型的快速发展逐渐渗透到数字流行病学建模和传播动力学研究中，在一定程度上重塑了传统研究范式。一方面，基于大语言模型的智能体运用自然语言模拟网络节点的属性、状态转变、交互等，取代由规则或统计方法驱动的传统智能体，从而提升传播模拟的复杂性和真实性。另一方面，大语言模型本身也可视为传播过程中的新型信息生产者和放大器。其大规模内容生成能力改变了信息传播的速度和规模等关键方面，因而成为影响信息传播的重要外部变量。以下各节将详细阐述这两个方面，并总结相关研究及其主要发现。

表5：基于大语言模型的流行病智能体建模方法总结。第三列记录了建模过程中使用的具体大语言模型。规模指参与建模过程的智能体数量。智能体的记忆分为三种类型：感觉记忆，用于感知；短期记忆，支持实时决策；长期记忆，有助于持续保持知识 [201]。网络结构指用于定义智能体如何相互连接和交互的拓扑形式，包括随机网络、无标度网络等。交互描述智能体之间的信息交换。同质交互仅发生在同类型智能体之间，而异质交互则涉及不同类型智能体之间的通信。

场景	研究	LLM	规模	记忆	网络结构	状态表示	交互
数字	Törnberg 等人 [178]	GPT-3.5	500	感官	-	文本	同质
	Li 等人 [62]	gpt-3.5-turbo-1106	288/300	感官	随机、无标度、高经纪性网络-work	Rule	同质
	Mou 等人 [69]	gpt-3.5-turbo	1K	长期	现实世界网络	文本、规则	异质
	FUSE [67]	gpt-4o-mini	40	短期、长期	随机，无标度，高聚类网-work	文本	异质
	Hu等人 [202]	gpt-4o-mini	100/168	感官	随机、无标度、小世界，real世界网络	文本	同质
	FDE-LLM [70]	ChatGLM-4	[206, 9890]	感官	-	文本、规则	异质
	Li 等人 [203]	-	-	-	-	Rule	同质
	Ju等人 [204]	Vicuna-7B，Llama-3-8B，Gemma-7B	[2, 10]	长期	-	文本	异质
	Zheng等人 [205]	GPT-4o	15	长期	小世界，规模-自由网络	文本	同质
	TrendSim [206]	GLM-3-turbo	1K	感官、短期、长期	-	文本、规则	同质
	Y Social [129]	Llama-3-7B	[1K，2K]	感官	-	文本	同质
	OASIS [179]	Llama-3-8B-Instruct	1M	短期、长期	无标度网络	文本、规则	同质
	LLM-AIDSim [130]	Llama 3	333	感官	现实世界网络	文本	同质
	MOSAIC [131]	GPT-4o	[50, 130]	长期	无标度网络	文本	异质
	Nasim 等人 [207]	gpt	-	-	-	文本、规则	异质
	S^3 [128]	GPT-3.5，ChatGLM-6B	8,563/17,945	长期	现实世界网络	文本	异质
	Nudo 等人 [89]	Gemini、Mistral、DeepSeek	1,186	感官	-	文本	同质
生物学	Williams等人 [61]	gpt-3.5-turbo-0301	100/1K	感官	-	文本、规则	同质
	Villaplana 等人 [80]	GPT-3.5	576	感官	-	文本	同质

3.1.1. 大语言模型作为数字流行病建模的增强工具

传统的数字流行病建模通常依赖预先设定的数学方程或规则来预测不同类型信息的传播趋势。尽管方法直观，但这类方法往往过度依赖对观点和消息的数值抽象，因而简化了现实场景的复杂性。因此，它们无法捕捉信息的语义内容，也不能充分反映人类推理、交流和决策的灵活性。如今，越来越多的研究人员开始利用大语言模型能力进行环境感知、逻辑推理和自主规划，以解决这些不足 [19]。在基于主体的建模框架中，研究者提出以大语言模型驱动的智能体替代传统的基于规则的智能体，从而模拟个体对信息扩散的行为响应，并更贴近人类认知 [62; 63; 64; 208]。

在早期尝试中，Li 等人 [62] 采用了简单的提示词，例如“你是一个 *[角色设定]*。你对这则*[新闻]*有什么看法？然后，根据你的人格，你愿意与朋友分享这则 *[新闻]* 吗？请回答‘是’或‘否’，并说明理由。”，以指示大语言模型推理预定义智能体是否会传播给定的错误信息。他们将人格视为控制变量，重点考察三种网络结构（随机网络、无标度网络和高经纪网络）如何影响假新闻的传播。实验结果表明，外向性和开放性得分较高的参与者更有可能传播信息 [209； 210]。在结构层面，由于高度连接节点的加速效应，无标度网络表现出最快的传播速度；而高经纪网络中稀疏的簇间连接形成了瓶颈，因此传播速度最慢。

然而，在现实社交网络中，个体是否采纳并传播信息，不仅受年龄、性别和人格影响，也受历史行为、社会认同和环境因素影响。为解决这一局限，Liu 等人提出了一种基于大语言模型的传播模拟框架（FPS），通过系统整合历史记忆、环境交互和身份特征，利用大语言模型模拟类人认知与行为 [63]。如图 6 所示，FPS 由两个核心模块组成：动态观点智能体（DOA）和智能体交互模拟器（AIS）。DOA 为每个智能体配备短期记忆，用于记录日常互动；长期记忆，用于捕捉更广泛的背景信息；以及反思机制，用于模拟人类思维过程并更新对新闻的态度。所有这些功能均通过大语言模型的提示词技术实现。另一方面，AIS 构建智能体进行交互的环境。如图 5 所示，FPS 中的观点交流通过自然语言交流进行，与现实世界类似，而非采用抽象的数值表示。研究结果发现，与恐怖主义、自然灾害、《科学》、民间传说或金融相关的错误信息相比，政治错误信息传播得明显更快。此外，研究还发现，宜人性和神经质得分较高的个体更容易相信并传播虚假信息。Wang等人 [65] 提出了一种社会模拟框架，其采用了与 FPS 几乎相同的基于大语言模型的智能体属性和传播行为设置。该框架引入了常用的推荐算法来推送相似观点，旨在研究回音室效应和意见极化现象。

与上述研究类似，Chuang等人 [66]同样采用大语言模型来减轻失真和过度-

You are simulating a real person with [*the profile: name, age, gender, personalities, education level, culture, etc.*]. Given your [*previous personal opinion*] / [*observations*] and the new information in your [*long memory*] / [*history behaviors*], update your opinion. Compose a tweet expressing your opinion. Use 0 for disbelief and 1 for belief to indicate your opinion. Provide reasoning behind your tweet and explain the rationale for your belief.

图5：基于大语言模型的智能体的观点推理提示模板示例。

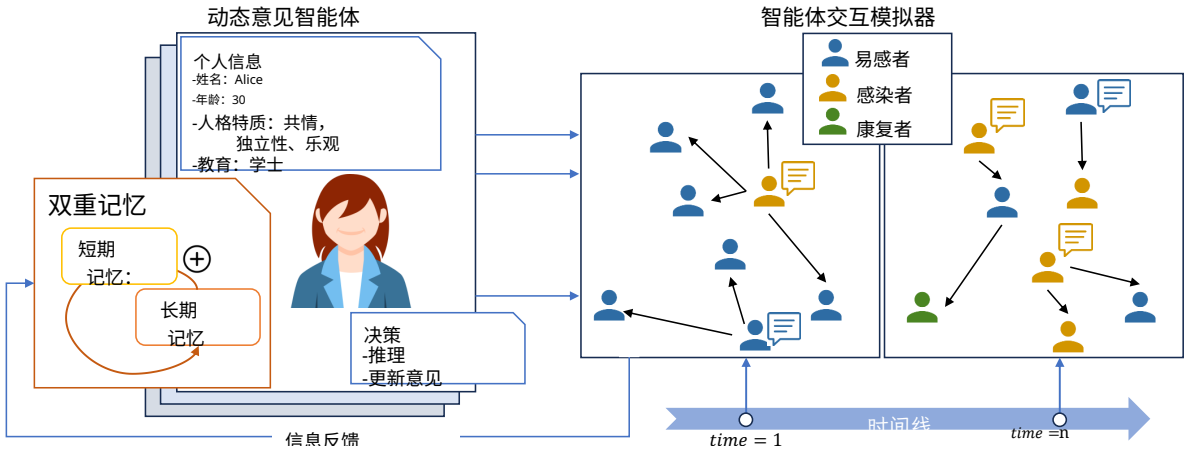


图 6：FPS框架示意图。该框架模拟数字疫情传播过程，由两个主要组件构成：动态观点智能体和智能体交互模拟器。动态观点智能体利用短期记忆记录与其他智能体的近期互动，利用长期记忆存储背景信息，并通过反思机制模拟类似人类的认知过程，更新其对特定信息的态度。智能体交互模拟器构建了一个交互式环境，智能体在其中通过自然语言交流彼此沟通，而非依赖预定义的参数或规则。来源：转载自参考文献 [63]。

传统基于主体的模型所固有的简化，这些模型依赖于对信息的数值抽象以及基于规则的互动。与以往方法不同，他们的工作将确认偏误纳入智能体的观点演化，并引入了一种独特的记忆更新机制，以考察这些因素如何塑造观点动力学。其模拟框架采用说话者-倾听者设定，基于大语言模型的智能体通过以一对一方式接收当前基于大语言模型的说话者所提供的信息来更新自身观点，而不是汇总其先前遇到的全部观点（一对多）。在该框架下，研究采用了累积记忆机制，每一次新的交互都按序列保留在记忆中，从而完整记录智能体观点演化的轨迹。通过为基于大语言模型的智能体赋予不同程度的确认偏误，研究发现，在没有确认偏误时，大语言模型驱动的智能体倾向于迅速收敛至科学共识。然而，一旦引入确认偏误，观点多样性便显著增加，复现了传统基于主体的模型中常见的极化现象，且偏见越强，多样性越大。与此同时，该研究揭示了用大语言模型构建智能体的一个局限：无论被赋予何种角色，基于大语言模型的智能体通常倾向于拒绝不准确的信息并认可事实性内容，这限制了它们在模拟持续抵制事实证据的个体时的真实性。值得注意的是，共识是指从社会主体之间的互动中涌现出的有序状态 [153]。Ashery等人 [211] 表明，由大语言模型控制的智能体能够通过彼此的互动自发形成共识。同样，De Marzo等人 [212] 发现，相互交互的智能体可以彼此施加影响力，从而使群体可能自发地就任意决策进行协调。相关研究为理解基于大语言模型的数字疫情传播提供了有价值的洞察力，进一步讨论可见参考文献 [213；214；215；216；217；218； 219]。

除两类说话者—听众型基于大语言模型的智能体外，还有一些研究利用大语言模型构建了更多样化的智能体角色，以更好地捕捉数字流行病传播中的个体异质性和行为复杂性。例如，FUSE框架 [67] 采用提示工程构建了四类基于大语言模型的智能体：传播者、评论者、核查者和旁观者。这些智能体分别在社交网络中错误信息的演化过程中执行信息扩散、观点解读、事实核查和被动观察。与此同时，Ma等人 [68]设计了一组基于大语言模型的智能体，将人口学差异（如年龄、性别、教育水平、政治倾向，以及对科学、政府和新闻的信任程度）与多样的生活经历相结合，作为虚拟参与者参与错误信息敏感性模拟测试。这些研究表明，大语言模型可用于生成多种智能体角色，从而在语义和行为层面再现复杂的社会传播现象。

然而，与传统的基于规则或基于概率的基于主体模型相比，基于大语言模型的智能体建模在计算效率和资源消耗方面存在劣势。经典基于主体建模方法能够在短时间内实现大规模人口传播模拟。例如，EpiSimdemics 能够在 NCSA 的 Blue Waters 超级计算系统（配备 352,000 个计算核心）上，于一段时期内完成数亿人的传播模拟

即 120 天仅需 12 秒 [220]；而 Epiabm 可以模拟直布罗陀地区（人群约 34,000）80 天内的感染传播，在 AMD 3600X 处理器（6 核，3.8 GHz）上该模拟仅耗时 8 秒 [221]。相比之下，基于大语言模型的智能体系统通常需要在每一轮交互或状态更新中调用大规模语言模型进行推断，其计算成本随智能体规模和交互轮次的增加而迅速上升。如表6所示，现有基于大语言模型的智能体系统在数千至数万的智能体规模下，通常需要数小时甚至数十小时的运行时间，并依赖高性能计算资源或高昂的推理成本。这种高计算开销限制了基于大语言模型的方法在大规模、长时间传播模拟中的可扩展性。

表6：基于大语言模型的智能体系统的计算成本。

方法	模型	规模	硬件	延迟	Cost
AgentSociety [90]	DeepSeek-V3	10K	华为云 c7.16xlarge.4	458.82 秒/轮次	-
GABM [61]	gpt-3.5-turbo	1K	32 GB 随机存取存储器中央处理器	80 小时/轮次	$20/轮次
TrendSim[206]	GLM-3-turbo	1K	英特尔至强 Gold-5118（48 核心）中央处理器	16h	-
OASIS [179]	Llama3-8b-instruct	1M	24 个 A100 GPU	2.8 小时/轮次	-

为应对大规模社交网络用户模拟在成本和效率方面的挑战，Mou 等人 [69]提出了一种混合仿真框架。该框架利用社交网络用户固有的帕累托分布特征，将其分为两类：核心用户（如意见领袖）和普通用户。针对这些角色采用不同的建模方法：使用大语言模型刻画并驱动核心用户，以捕捉人类复杂的行为模式；而数量庞大的普通用户则由传统的基于主体的模型控制，从而实现更高效的大规模用户模拟。

同样，融合动力学方程—大语言模型（FDE-LLM）框架 [70]也将社交网络用户划分为意见领袖和追随者，前者通过基于大语言模型的角色扮演进行建模，后者则使用传统的基于主体的模型进行模拟。为在大语言模型的引导下维持群体态度的自然衰退，并提高模拟群体内意见传播的效率，FDE-LLM 框架融合了元胞自动机（CA）和 SIR模型的约束。在这一设计中，意见领袖通过大语言模型驱动的角色扮演来实现，其观点变化受元胞自动机约束；追随者则嵌入结合元胞自动机与 SIR模型的动态系统中。实验结果表明，这一设计显著提升了模拟的准确性和预测效率，尤其是在涉及反转新闻的场景中。

在上述两项研究中，核心用户与普通用户的划分通常是固定的，缺乏动态适应能力，因而限制了传播模拟的灵活性。为解决这一局限，Liu 等人 [71]提出了一种动态分层社交网络仿真框架，能够对数百万个体进行建模。该框架引入了一种基于信息茧房理论的多智能体动态交互策略，可根据每个时间步的信息混乱程度，自适应地识别核心智能体和普通智能体。同样，核心智能体由大语言模型驱动，以捕捉复杂的推理过程；普通智能体则由基于主体的模型控制，从而实现大规模可扩展性。核心智能体的数量及智能体之间的交互模式都会随时间动态调整，进而增强模拟的灵活性和真实感。此外，该框架还纳入了分层协作网络，其中优先连接有助于形成意见领袖，三角连接则促进紧密本地社群的形成，从而加速信息扩散。实验结果表明，与现有模型相比，该方法在模拟现实世界谣言传播时具有更高的一致性和准确性，平均观点偏差降低了 64%；研究还进一步发现，社交网络中紧密连接的本地社群结构是推动谣言快速传播的关键因素之一。

除了建模效率之外，生成逼真且可扩展的社交网络对于数字流行病建模也至关重要，因为网络结构在很大程度上决定了传播路径和规模。遗憾的是，传统的基于规则的模型难以捕捉现实世界交互的复杂性，而深度学习方法往往依赖大规模标注数据集。为克服这些局限，Ji 等人 [64] 提出了零样本社交网络生成框架 *GraphAgentGenerator*（*GAG*）。通过将由大语言模型驱动的多智能体机制引入图生成过程，GAG 使智能体能够通过模拟交互动态生成节点和边，从而生成更接近现实的社交图。GAG 框架包含两类节点：(i) 行动者节点，以基于大语言模型的智能体实例化；(ii) 表示内容的项目节点。边通过智能体与内容的交互逐步形成。对于每个行动者，其用户档案由大语言模型生成，同时智能体会维护一个记录其活动历史的文本记忆组件。在每个仿真轮次中，智能体在设计好的动作提示引导下，依据历史记忆和上下文检索采取行动，并通过反思更新状态。GAG 中智能体的活动遵循帕累托分布，约 20% 的高活跃度智能体产生了大部分交互。实验结果表明，GAG 生成的社交网络具有现实系统中观察到的七种宏观属性，包括幂律度分布、小世界特征和收缩直径。此外，该框架展现出很强的可扩展性，支持生成多种网络类型，规模最高可达 10^5 个节点和 10^7 条边。这项工作表明，将大语言模型融入社交网络生成，不仅能够在语义层面模拟个体交互，还能复现统计

物理学中已充分确立的复杂网络规律。因此，GAG 为数字流行病传播建模提供了一个逼真且可扩展的网络环境。

Cheng 等人 [72] 提出了一种社交网络仿真系统，将大语言模型驱动的智能体与交互式可视分析相结合。该系统将建模与可视化技术相结合，提出了一种交互式方法，用于建模和研究数字流行病，使研究人员能够在复杂、动态的场景中开展迭代探索与分析。该系统展示了具有不同特征的大语言模型驱动的智能体在处理数字信息、调节情绪和立场、提供反馈以及影响事件演变方面的差异，从而以可视化方式揭示社交网络舆情事件的动态机制。

除了上述受基于主体建模启发的建模方法外，其他研究还将大语言模型集成到不同建模框架的特定模块中，以增强其功能。例如，Jiang 等人 [73] 利用大语言模型推断传播树中各节点对帖子的立场，从而为节点分配状态标签，并促进流行病学信息表示学习优化目标的实现。与此同时，BotSim [74] 系统利用大语言模型构建用户节点，使其在数字流行病中生成和传播有害信息，并进一步使用定制提示词指示大语言模型完成任务分解、感知真实社交网络环境等任务。

3.1.2. 大语言模型作为数字疫情传播中的影响因素

在对数字疫情传播动力学进行建模时，大语言模型的出现引入了一系列影响传播格局的新型外部因素。这些因素包括：大语言模型作为数字流行病的信息源；由大语言模型控制的智能体参与以人类为中心的传播过程；以及由大语言模型驱动的智能体网络中发生的传播。随后，我们将详细分析各个场景，并讨论这些外部因素如何共同影响数字疫情传播的建模。

首先，大语言模型成为新的内容生产者，从而扩展了“信息源”的范围 [222；223；224； 99]。它们能够以极低的成本大规模生成几乎与人类写作难以区分的内容 [225；43； 226], ，并可能因幻觉而无意中生成错误信息，也可能在被利用来编造具有说服性和误导性的叙事时有意生成错误信息 [227； 44]。从接种理论的角度来看 [228], ，这类由大语言模型生成的输出可被视为新的“病原体”，它们不断演变、具有适应性且难以检测 [223]。因此，大语言模型不仅增加了此类“病原体”的流通量，还通过增强其多样性和可信度，加速错误信息扩散 [227； 44]。从动力学角度看，大语言模型持续生成内容相当于一个外源注入项，提高了错误信息的再生数及其有效感染率，从而可能增强系统层面的传播能力，并改变暴发阈值和稳态。

其次，大语言模型控制的智能体可能参与以人为中心的数字流行病传播，与人类用户互动，并影响信息、观点或错误信息的扩散。这种参与会产生两种影响。一方面，大语言模型会在训练过程中内化并放大社会偏见，使其生成的输出天然带有意识形态倾向或两极分化倾向 [229；230； 231]。由于许多开放式问答系统和交互平台依赖大语言模型生成响应，这些带有偏见的输出可能直接影响用户，并通过社交网络上的转发机制得到进一步放大，从而传播特定叙事。这意味着，个体的易感性和传播概率不再是同质的，而是受到偏见的异质调节。另一方面，由大语言模型驱动的社交机器人可以部署于在线网络中，模拟类似人类的互动，并通过定向互动和个性化内容注入改变扩散路径 [232；233； 44]。近期研究表明，此类机器人甚至能够以很高的准确性推断用户属性，并实施定向干预，作为具有智能干预策略的“超级传播节点”发挥作用 [234]。Hitz 等人 [235]发现，与人类受试者相比，由大语言模型驱动的智能体往往具有更低的采纳阈值，因此可能充当放大器，导致更快速、更广泛的数字疫情传播。

第三，传播动力学也可能产生于由大量大语言模型驱动的智能体所构成的网络之中；这些智能体可能受不同的大语言模型控制，人工代理之间的互动本身便会催生复杂的传播模式。此类传播过程的动力学及其特征本身即构成一个新兴的研究课题。这一探究方向属于新兴的跨学科领域——机器行为；正如Rahwan 等人 [236], 所阐述的，该领域呼吁对智能机器开展科学研究，不是将其视为工程制品，而是将其视为一类具有特征性行为模式与互动生态的行动者。近期研究已开始从实证上探索这一方向。例如，Marzo 等人 [75] 考察了使用GPT-3.5-turbo实现的多个生成式智能体之间的互动，分析信息与行为如何在完全由人工代理组成的人群中传播。他们发现，此类智能体不仅能够模仿个体的人类语言行为，还能够展现出人类社会所固有的集体现象。类似研究可见参考文献 [76； 77]。这些研究有助于在更广泛的机器行为语境下，更好地理解大语言模型增强型智能体作为机器所表现出的传播行为与传播模式。与此同时，此类智能体—智能体传播过程并非与人类参与相互隔绝。在网络空间中由人工代理交换并放大的信息，最终可能被仍嵌入更广泛社会技术系统中的人类用户所观察、消费或进一步传播。因此，理解大语言模型驱动的智能体之间的传播动力学，可为机器生成的信息如何与以人为中心的数字疫情过程相互作用并反馈至其中，提供重要洞察。

上述讨论的三个因素都会影响数字疫情传播。在实践中，它们并非彼此孤立，而可能同时存在于网络空间的复杂系统中。例如，Lu 等人 [78] 研究了由大语言模型驱动的

对抗性社会影响在在线信息扩散中的作用。他们的研究指出，随着生成式人工智能的发展，大语言模型能够生成极其逼真、近似人类撰写的文本，从而显著提升社交机器人的效果。因此，由大语言模型驱动的机器人账号不仅能批量生成与新闻真实性相悖的评论，还能针对人类用户的帖子生成对抗性回复，从而营造出对立观点占主导地位的虚假印象。在一项涉及 176 名参与者的随机对照实验中，作者发现，这种由大语言模型驱动的对抗性影响削弱了个体区分真实新闻与虚假新闻的能力，也降低了他们分享真实信息的意愿。值得注意的是，与对抗性回复相比，直接针对新闻真实性发布的对抗性评论在抑制真实新闻传播方面更为有效。此外，这种影响受到参与者政治倾向的调节：当新闻内容与其政治立场相冲突时，由大语言模型驱动的对抗性影响会显著加剧误判。这些发现凸显了大语言模型在信息扩散过程中的双重作用：它们不仅生成误导性内容，还通过对抗性社会互动扰乱人类认知与行为，从而加剧错误信息的大规模传播。从传播动力学的角度来看，这项研究提供了重要洞察力。首先，由大语言模型驱动的社会影响可视为一种外部动力学因素，它并非中性的随机噪声，而是一种系统性扰动，通过“评论”和“回复”在局部节点传播对抗性观点，从而有效地在扩散模型中引入定向扰动场。其次，由于对抗性影响会降低分享真实信息的意愿，同时提高虚假信息的有效传播率，因此它可能降低信息扩散的临界阈值，使错误信息疫情更有可能越过临界点。受物理学启发的“数字流行病学”和复杂系统的传统方法通常假设扰动是均匀或随机的，而由大语言模型驱动的扰动则具有智能性、方向性和动态适应性。这表明，未来的传播模型需要明确纳入基于大语言模型的机器人账号作为主动传播者所产生的动力学耦合，并考虑此类定向操纵如何改变系统稳定性与临界性。

此外，Bandara [79] 研究了大语言模型生成的幻觉如何成为虚假信息的来源，以及它们如何系统性地放大阴谋论和假新闻的传播。简言之，大语言模型幻觉通过三个层次逐步发展：始于模型架构的技术缺陷，继而生成虚假内容，最终演变为社会传播。这使大语言模型幻觉成为新型虚假信息生产与传播的核心驱动力。为解释大语言模型如何助长虚假信息生态系统的放大效应，该研究将其影响归纳为三个主要方面。(a) 阴谋论传播：大语言模型将无关事件联系起来，构建虚假的关联，将其描述为所谓的“隐秘阴谋”，从而制造证据确凿的假象。在封闭社群中，人们对这类内容的集体解读，加上算法推荐机制，会强化这些幻觉，使其成为群体共享的信念。此外，当幻觉内容将真实历史事件与当代谣言相结合时，便会出现历史嫁接，从而提高虚假信息的表面可信度。(b) 公共领域的衰退：算法推荐和回音室效应加剧了意识形态极化，使用户陷入不断强化既有偏见的反馈回路。与此同时，随着幻觉内容削弱主流媒体和学术机构在人们心中的权威，制度信任也随之崩塌。信任的恶化进一步带来民主风险：公众无法依据真实信息形成共识，进而妨碍理性政策制定。(c) 大规模假新闻生产：借助自动化内容流水线，与编排工具集成的大语言模型能够协调开展跨平台的虚假新闻传播。流量驱动优化会依据实时互动指标调整生成内容，优先推送情绪煽动性强、立场极端的叙事。同时，通过关键词优化操纵搜索引擎，可使虚假信息占据搜索结果中的显著位置。总之，大语言模型的集成不仅催生了新的病原体，还作为新的传播媒介深度介入网络动态，为数字疫情传播建模带来了新的挑战。

3.2. 基于大语言模型的生物疫情建模

生物疫情建模往往与预测、干预等下游任务紧密结合，旨在刻画潜在的传播动力学，并支持流行病学预测与管理策略。随着大规模数据和跨学科技术日益可得，研究人员已开始探索大语言模型如何增强这一过程的智能化与自动化，因为其在上下文感知推断、跨源语义整合和潜在模式抽象方面的强大能力，为理解疾病传播的复杂动力学提供了颇具前景的基础 [237]。总体而言，现有研究可分为两类。第一类是让大语言模型直接参与建模过程，使其在传播建模、动态推理或预测中发挥积极作用。第二类侧重于将大语言模型间接用作辅助工具，以协助代码生成、模型优化和数据拟合等任务。前者凸显大语言模型的推理与机制性能力，后者则强调其提升建模效率与自动化水平的潜力。以下各小节将详细阐述这两条研究路线。

3.2.1. 大语言模型直接参与建模过程

与数字流行病中基于大语言模型的基于主体建模方法类似，一些生物疫情建模研究也使用基于大语言模型的智能体直接模拟流行病动态 [80]。例如，Villaplana 等人 [80] 使用 ChatGPT-3.5 构建了一组具有异质属性和信息背景的智能体。通过提示工程，他们为每个智能体设定了不同的行为特征，以模拟人类决策过程，从而弥补传统 SIR 模型缺乏行为反馈机制的不足。在实验中，他们探讨了

关于病原体特征及其对生物流行病传播影响的两个关键问题：智能体每天与多少人互动，以及这些互动中是否保持社交距离。结果表明，随着智能体接收更多有关流行病的信息，病例增长曲线可能逐渐趋于平缓，峰值变得不那么明显，甚至消失，这表明信息传播能够显著影响疾病传播动态。

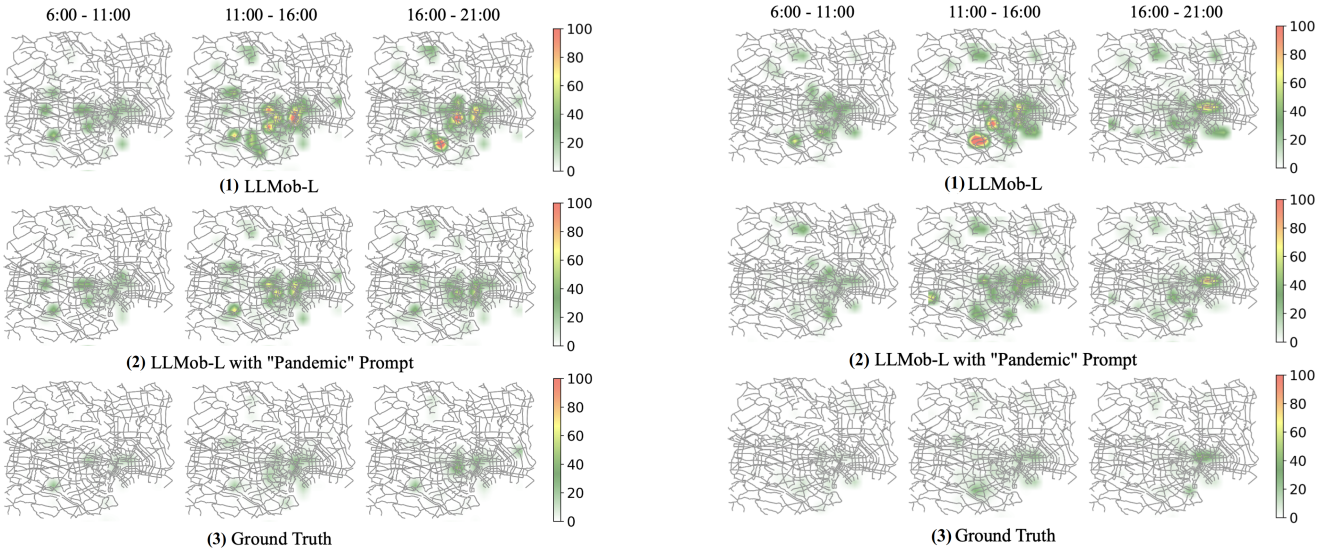
除了基于大语言模型的智能体建模范式外，还有一些研究使用大语言模型直接建模传染病传播的时间演变。EpiLLM框架 [81]将大语言模型用作时空疫情建模的基础模型，并采用自回归方法刻画传播过程。传播过程被划分为多个时间片段，在此期间，图神经网络将流行病学特征（例如感染人数）编码为词元化输入，供大语言模型使用。随后，模型根据前 N 个N -1 片段预测第N个片段的传播结果，从而实现对时空疫情演变的生成式建模。

此外，Moon等人 [82] 和Du等人 [83] 提出了基于大语言模型的多模态疫情建模框架，整合疫情统计数据、公共卫生政策文本和时间序列等多种上下文信息，以支持更准确的预测。类似地，Kang等人 [84]利用大语言模型学习疫情建模涉及的多模态信息，丰富患者、疾病和出院摘要的语义表示，从而构建更细致、更高阶的健康状态表示矩阵。与众不同的是，他们通过构建疾病共现网络，提出了一种由大语言模型增强的患者内多模态建模策略，以便更好地理解单个患者体内疾病的进展与相互作用。

3.2.2. 作为辅助工具间接参与流行病建模的大语言模型

除了让大语言模型直接参与流行病建模外，一些研究还将其用作辅助工具，以间接支持建模过程。例如，Kwok 等人 [85] 提出使用 ChatGPT 协助开发传染病传播模型。通过自然语言交互，ChatGPT 执行代码生成、优化和调试任务，使研究人员能够构建 SEIR模型，并将其拟合到流行病学数据。该方法显著降低了技术门槛，使非编程领域的专家也能快速构建传播模型。

此外，大语言模型还被应用于模型建模之后，根据建模信息提供判断和决策功能。Zaslavsky 等人[86] 将大语言模型与具备空间感知能力的基于智能体的模型和系统动力学（SD）模型相结合，利用这些模型的空间推理和知识表示能力，增强其对人口分布模式的敏感性。例如，在 COVID-19传播模拟中，大语言模型为参与者生成具有空间偏好的行为，并根据参与者的健康状况和环境中的病毒载量动态调整接触概率，从而使传播过程更符合真实的社会空间交互特征。



(a) 艺术与娱乐。　　*(b)* 专业及其他场所。

图 7：“艺术与娱乐”和“专业及其他场所”活动在疫情情景下的活动热力图。“LLMob-L”表示集成了学习驱动的动机检索机制的 LLMob 框架变体。使用“大流行病”提示，使大语言模型智能体能够规划疫情期间的活动，具体如下：“现在正值大流行病期间。政府已要求居民推迟出行和活动，并尽可能远程办公”。来源：改编自参考文献 [87]

基于空间推理与行为适应的这种融合，研究人员进一步将大语言模型应用于人类移动预测，而这是流行病建模的重要组成部分，因为传染病的传播在很大程度上取决于个体和群体的空间分布及其互动。因此，准确模拟和预测人类移动模式已成为提升生物流行病模型保真度的关键。近年来，一些

研究提出了专用于高保真人类移动建模与预测的基于大语言模型的框架，为疫情传播建模注入了新的动力 [87]。例如，Wang等人提出了个人移动生成大语言模型智能体框架（LLMob），该框架将城市环境中的个体视为基于大语言模型的智能体，通过联合建模习惯性活动模式和情境驱动动机，生成受社会因素约束且时空一致的移动轨迹 [87； 238]。研究发现，如图 7 所示，该模型在疫情提示下能够再现更真实的时空活动分布。这种更高的真实感源于模型纳入了有关疫情影响和政府应对的先验知识，使大语言模型智能体能够采取符合现实世界行为适应的行动。

4. 流行病检测与监测：大语言模型增强的传播动力学感知

流行病传播过程的检测与监测，主要用于感知现实世界中传染物的对象、行为及演化特征，包括识别传播 [239的源头或起源； 240], 从网络结构中识别早期预警信号 [241]，并监测疾病或信息扩散的当前趋势 [242；243；244]。从物理学家的角度来看，传播动力学主要指对传染物如何通过复杂网络传播进行理论建模、模拟和分析 [149]。因此，通过感知传播过程，检测与监测能够为传播动力学模型的理论表征与参数校准提供现实依据，从而在传播动力学的理论研究与实际应用之间发挥关键的桥梁作用。

本节探讨大语言模型在感知传播过程涌现特征及当前趋势方面的作用。通过分析海量文本与多模态数据，大语言模型能够检测弱信号、追踪暴发源头并表征传播状态。具体而言，本节综述大语言模型如何通过识别谣言或错误信息的早期信号来增强对数字流行病的感知，以及它们如何协助开展生物疫情监测，以实现早期预警、症状信号检测和持续趋势监测。

4.1. 大语言模型增强的数字疫情感知

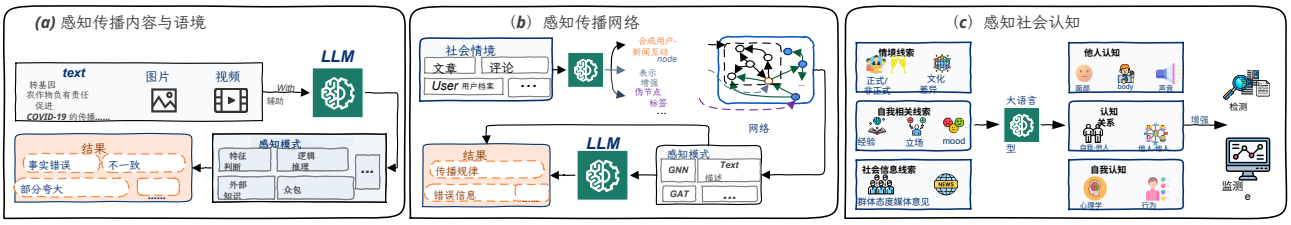


图8：数字流行病中的不同感知目标。

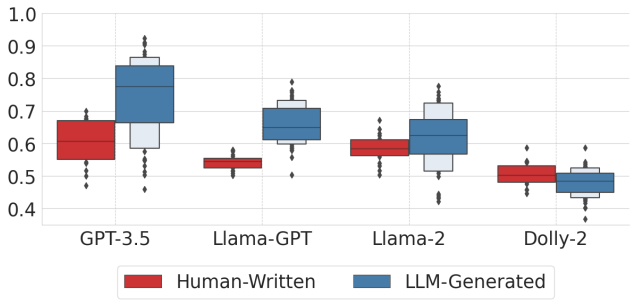
凭借在语言理解、跨模态信息处理和语境推理方面的综合能力，大语言模型日益应用于数字流行病传播的检测与监测。大语言模型不仅能够识别和解读复杂且快速演变的在线环境中的传播内容，还能整合来自异构平台和多模态的信息，以揭示传播路径、语义关联和潜在动态特征，为感知数字疫情传播过程提供新的技术手段。总体而言，大语言模型带来的数字流行病感知增强体现在多个维度。一方面，大语言模型可以通过考察叙事逻辑、立场倾向和背景知识，在内容和语境层面对信息传播进行细粒度分析，从而协助识别可能存在的虚假、误导性或操纵性内容。另一方面，可以利用大语言模型的生成和推理能力，对传播网络的结构进行补充、重建和推断，从而表征不同信息单元之间的传播路径、结构模式及其演变过程。此外，大语言模型还为感知社会认知提供了可行途径。它们能够从大规模文本与多模态数据中捕捉与传播相关的细微社会认知线索，例如情绪表达的变化、立场演变和互动语境，从而将社会认知状态的表征引入传播感知，使其超越传统的结构与行为感知。

根据不同的感知目标，现有研究大致可分为三个方向：(a) 感知传播内容和语境，重点是利用大语言模型的理解与推理能力，分析文本语义、基于评论的上下文和外部知识，以揭示内容层面的造假与扭曲；(b) 感知传播网络，重点是利用大语言模型的生成与结构推断能力，重建、恢复和分析信息传播的路径、模式及演化过程；(c) 感知社会认知，重点是利用大语言模型捕捉情绪、立场及其在传播过程中的变化等认知线索，描绘个体对他人、关系和自我的认知，以弥补传统感知在主观认知因素方面的不足。图 8 展示了感知上述目标的常见范式。下文各小节将对此进行详细阐述。

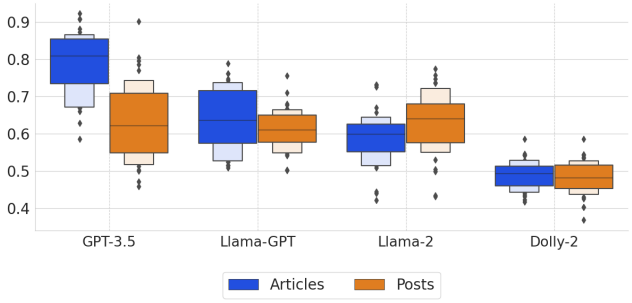
4.1.1. 感知传播内容与语境

有关感知传播语义和语境的研究，关注数字流行病内容的语义与语境层面，旨在利用大语言模型的理解与推理能力，检测错误信息、谣言、操纵性叙事等中的潜在语义偏差与真实性问题 [93; 94; 95; 96; 97; 98]。总体而言，这些研究可分为三个主要方向：利用大语言模型解读主题、立场和话语语义；借助外部知识和基于评论的信息提升上下文感知；以及构建基于大语言模型的智能体框架，模拟人类辩论以推理真实性。

首先，研究人员利用大语言模型解读传播内容的主题、立场和话语语义，从而捕捉数字流行病背后的隐含语义线索。谣言、假新闻和病毒式传播帖子等现象，往往依托于包含事实失真、夸张叙事或有害内容的文本载体 [245; 246; 247]。Lucas等人 [99] 研究了大语言模型的零样本错误信息检测能力，考察其能否在无需额外训练的情况下，通过上下文推理实现自我检测。如图9所示，他们发现，与机器生成的错误信息相比，大语言模型更难识别人类撰写的错误信息。在输入长度方面，大语言模型对新闻文章（长输入）的零样本检测性能也优于对社交网络帖子（短输入）的检测性能。这些发现表明，大语言模型的语境推理机制在处理信息丰富、结构良好的文本时更为稳健，但在感知零散或噪声较多的社交网络语境时仍存在局限。Ma等人 [100]提出了基于大语言模型增强语义挖掘的假新闻检测（LESS4FD）方法。该方法通过专门设计的提示词引导大语言模型提取并嵌入新闻主题和真实实体，构建异质新闻—实体—主题图，并采用带有一致性学习准则的广义 PageRank 模型传播自适应特征。这使模型能够联合捕捉每篇新闻周围的局部语义以及所有相关新闻之间的全局语义，从而识别单篇新闻文本与更广泛知识语境之间的不一致。与 TextCNN、TextGCN、HAN、BERT 和 Sentence-BERT 等传统分类器相比，LESS4FD 在 MM COVID 数据集上的准确性提升超过17.5%，F1 分数提升超过16.1% [100] 另一方面，一些研究将大语言模型用作内容分析器，以获取立场和主题等信息，而非直接将其用作虚假信息检测器。例如，Yang等人 [101] 开发了大语言模型增强的多实例学习（ LLM-MIL ）框架，将大语言模型生成的立场解释融入层次化立场树注意力机制，以提升谣言检测的可解释性和准确性。此外，Krykoniuk 等人 [102] 通过在话语层面比较大语言模型预测的正常话语语言轨迹与真实用户回应，检测主题“偏离”。上述研究主要关注基于文本或图像的虚假信息，对基于视频的假新闻检测关注相对较少。为弥补这一不足，Hu等人 [103] 提出了大语言模型专业知识引导的多模态自适应融合（MAGE-fend）框架。该框架利用大语言模型生成视频的图像描述，为图像提供高层次语义信息。此外，MAGE-fend 利用大语言模型整合视频标题、字幕和封面图像描述中的信息，生成解释性推理文本，以评估新闻真实性并提供关键判断。与 Hou-SVM、FANVM、TikTec 和 SV-FEND 等不使用大语言模型的方法相比，MAGE-fend 在 TikCron 数据集上的准确性和 F1 分数均提升超过6.8% [103]



(a) 大语言模型对人类撰写的虚假信息与大语言模型生成的虚假信息进行零样本检测的性能。



(b) 大语言模型对（长篇）新闻文章与（短篇）社交网络帖子进行零样本虚假信息检测的性能。

图9：大语言模型零样本虚假信息检测性能分析。在子图 (a) 中，纵轴表示准确性；在子图 (b) 中，纵轴表示 F1分数。(a) 显示，大语言模型检测虚假信息的准确性高于检测人类撰写的虚假信息的准确性。(b) 表明，大多数大语言模型在文章上的虚假信息检测性能更强。来源：根据参考文献 [99] 转载

在语义理解的基础上，后续研究通过整合背景知识、立场解释和基于评论的上下文，增强了上下文感知能力，并发现这些外部相关信息能够有效支持模型对数字流行病的检测。Hu等人 [93] 指出，由于推理能力有限，小型语言模型（SLM）检测器仍难以感知多源线索和背景知识。因此，他们引入了一种范式，由大语言模型（LLM）充当“顾问”，提供多视角指导，并在假新闻检测过程中增强语境推理。此外，用户评论是评估错误信息的重要上下文信号，但现实世界数据集往往由于曝光偏差或用户沉默而存在评论稀疏问题。为弥补这一缺陷，Nan等人 [94] 提出将大语言模型用作“用户模拟器”和“评论生成器”。通过向大语言模型提供多样化的用户画像提示，并汇总来自多个子群体的生成评论，其方法

模拟了丰富多样的评论语境，提升了模型对假新闻的语义理解能力。类似地，Chen等人 [104] 也致力于通过提升用户评论或公众意见的多样性来增强假新闻检测。为充分利用公众意见，Chen等人 [104] 开发了双视角推理模块，利用大语言模型同时生成支持性和反驳性证据，从正面和负面两个视角分析包括图像、文本和用户回应在内的多模态内容，确保证据的全面性。在外部事实核查方面，Bai 等人 [105] 提出了全上下文检索与核验（FCRV）框架，将基于大语言模型的主张与实体抽取同检索增强生成（RAG）机制相结合。该模型在线检索相关证据，并综合内部上下文信息与外部事实，以评估传播内容的真实性，从而提高检测的可靠性。此外，Zeng等人 [106] 提出了人机混合推理框架，整合“模型优先”“人工工作者优先”和“元投票”策略，将模型预测与众包评估相结合，从而在复杂社会情境下实现更稳健的真实性评估。

最后，研究人员探索了基于大语言模型的智能体系统，以模拟人类辩论和推理，用于真实性评估。Liu 等人 [107] 提出了辩论增强真相（TED）框架，其中大语言模型充当辩论智能体，进行交叉质询、反驳和总结，以模拟人类认知推理并改进真实性判断。Li 等人 [108] 提出了 FactAgent，使大语言模型能够模仿人类专家核验新闻主张的行为。该系统将新闻核验分解为多步推理工作流，使大语言模型能够通过内部知识或外部工具，按序核验子主张并综合得出最终判定。为应对多模态假新闻检测中开放域知识获取和理解能力有限的挑战，Xie等人 [109] 实现了两个用于整合开放域知识的基于大语言模型的智能体：查询生成智能体通过从文本和图像中提取关键信息来生成知识检索查询；知识过滤智能体则对检索结果进行筛选，保留有助于提高准确性的关键信息。这两个智能体通过反馈进行协作，以增强整体可靠性，并在来自微博和 Twitter 社交网络的数据上实现了超过 83% 的检测准确性。

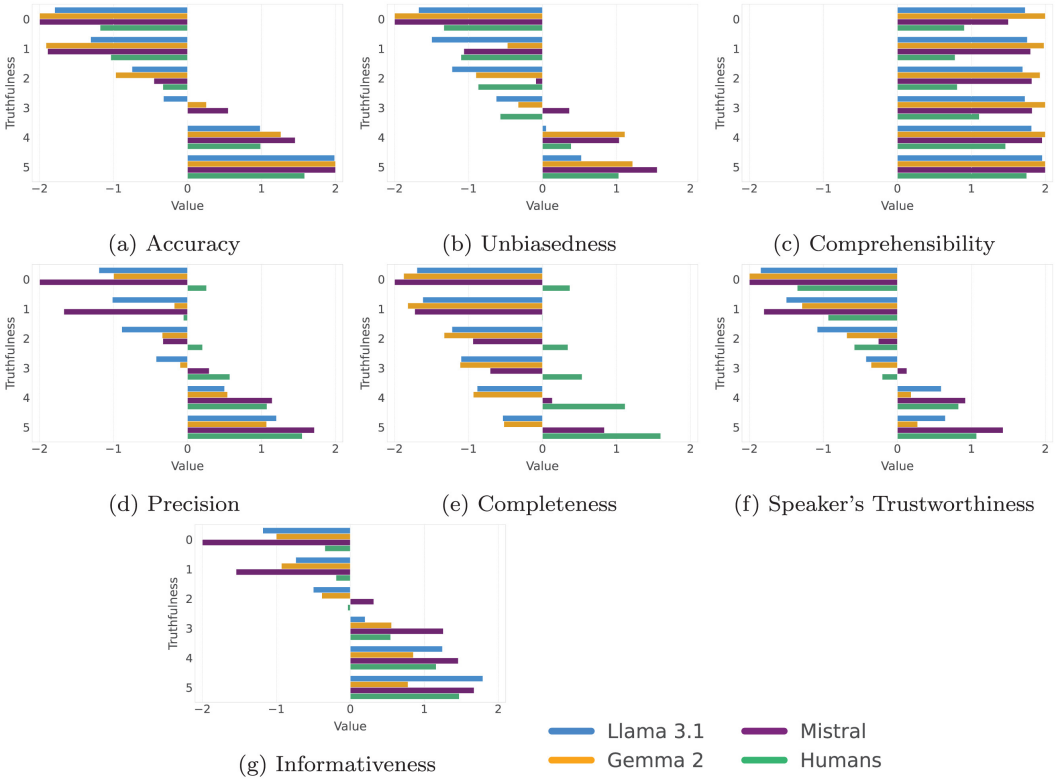


图10：在不同真实性水平下，基于大语言模型的智能体和人工评估者在各质量维度上的事实核查任务平均得分。纵轴表示真实性得分，横轴表示副标题中所示不同指标的得分。来源：转载自参考文献[110]

此外，Costabile 等人 [110] 探讨了在事实核查任务中使用基于大语言模型的智能体替代人类群体的可行性。他们构建了一群依据人口统计和意识形态特征（如性别、年龄、政治倾向）建模的大语言模型驱动的智能体，以模拟人类标注员，并要求这些智能体通过两个序列阶段评估主张的真实性：“证据选择”和“评估问卷填写”。评估过程中，每个智能体都会选择外部资源作为证据，并从多个维度评估主张，包括准确性、公平性、可信度等。图10展示了大语言模型智能体和人工评估者在各质量维度上的平均得分，揭示出两组之间存在明显不同的评估倾向。结果表明，大语言模型智能体倾向于给出更加两极化的判断，而人工评估者则偏好温和或中性的评分，且表现出更大的差异性，反映了更广泛的主观解读。最后，研究得出结论，生成式智能体表现出

其内部一致性显著高于人类标注员，并且能够有效利用外部证据评估主张真实性，其性能往往可与人类群体相当，甚至超过人类群体。这些发现表明，基于大语言模型的智能体有助于缓解错误信息治理中一致性和可扩展性方面的挑战。

4.1.2. 感知传播网络

在数字流行病中，信息传播不仅由消息内容决定，也取决于社交网络中的传播路径、结构模式和动态演化。传播网络刻画了信息如何被接收、回应和转发，揭示群体共识的形成、情绪放大，以及错误信息不断升级的机制 [248；249；250； 251]。然而，现实世界的用户交互数据往往不完整、零散，或难以获取，因此，假定图结构静态或能够观测完整网络的传统模型适用性较低。因此，近期研究利用大语言模型的生成、推理和结构抽象能力，重建、恢复并推断传播结构，从而支持数字流行病的检测与监测。

首先，为了应对网络缺失或仅能部分观测的问题，Wan等人 [111]利用大语言模型生成与新闻相关的多视角评论和用户反应，构建合成用户—新闻交互网络。他们进一步利用大语言模型推导出的隐式背景信号（例如立场、情感、外部知识）丰富节点表示，随后使用图神经网络在该网络上进行信息传播。

其次，为缓解数字流行病中的叙事碎片化和认知视角分歧，Li 等人 [112]提出了叙事重排序与视角融合（NRPF）方法。通过使用大语言模型进行多轮视角细化和叙事重排序，NRPF 汇总了分散在评论、引文和传播路径中的语义线索，从而有效描绘传播过程中的认知演化，并捕捉谣言传播模式及其结构。

此外，鉴于大语言模型在中等规模的结构化语境中推理效果更佳，Zeng等人 [113]对谣言传播网络进行剪枝，并将保留的核心节点转换为适用于大语言模型的文本表示。随后，他们设计了线索链提示，引导大语言模型逐步分析局部评论、全局评论和传播模式，最终生成谣言标签及可解释的推理输出。这缓解了大语言模型处理大规模结构化网络时能力不足的问题。

基于相连节点更有可能共享相同标签这一假设，Hu等人 [114] 提出了一种基于大语言模型的全局标签传播网络，以提高多模态错误信息检测的准确性。他们构建了一个跨模态新闻网络，其中每个新闻条目都被建模为一个节点，当模态内或模态间相似度超过阈值时，便在节点间建立边。为缓解有标签样本有限的问题，他们利用大语言模型为无标签节点生成高置信度伪标签，从而提供更多监督信号并稳定传播过程。

4.1.3. 感知社会认知

在现实世界的社会传播过程中，个体的行为和状态转变不仅受外部联系关系或传播结构的影响，也受其对自我、他人及人际关系的社会认知影响 [252]。个体的社会认知受到多种非结构化因素的影响，例如当前注意的内容、过往经历和社会背景 [253]。因此，即使处于相同的社交网络结构中，不同个体也可能对传播信息形成不同的解读，并对社会环境作出不同反应 [254；255]。

传统的传播感知方法，以及大多数关于信念演化和集体行为的社会动力学模型，往往强调可观测外部行为和结构特征——例如接触关系、传播路径或节点状态变化——并将个体简化为理性单元或规则驱动单元，认为它们会在交互情境中对外部刺激作出响应。这类假设往往忽视个体的内部心理过程和社会认知机制 [254]。计算社会科学领域近期关于社会感知的研究表明，对社交网络的主观表示能够提升社会动力学模型的描述效度和预测能力 [254]。不同于基于结构或行为的“客观”测量，这些主观表示捕捉了个体实际体验社会关系的方式，并受到情感态度、社会认同以及交互依赖度的共同影响。因此，将此类社会认知因素纳入传播感知与建模框架，可以提升社会动力学模型描述和预测现实世界传播过程的能力。

基于这一视角，大语言模型为感知传播过程中的社会认知提供了一条可扩展的技术路径。作为具备语言理解、语境推理和跨模态处理能力的感知引擎，大语言模型能够从与传播相关的大规模文本与多模态数据中提取认知相关线索，例如情绪表达和立场倾向的变化，以及它们在传播过程中的动态演化 [101；256；257； 258]。通过整合来自不同平台和不同时间尺度的异质观测，大语言模型能够对传播过程中的社会认知状态进行建模，超越传统的基于结构和行为的感知，从而为检测和监测任务提供更丰富、更真实的感知基础，并为传播动力学模型的表征和校准提供有价值的补充信息。

4.2. 大语言模型增强的生物流行病感知

鉴于传染病对公共卫生构成的重大威胁，传染病检测 [259] 和监测 [260; 261; 262； 263] 的主要目标是描述当前疾病负担和流行病学模式

，监测时间趋势，并识别疫情暴发和新兴病原体 [264]。这些工作为评估公共卫生风险的紧迫性，以及制定有效的应对策略和干预政策提供了关键支持，从而提高了在早期阶段遏制疫情暴发的可能性 [265]。根据世界卫生组织（WHO）网页上的定义，公共卫生监测是对健康相关数据进行持续、系统的收集、分析和解读[2]。这类数据已被直接用于为政策决策提供依据，并向公众及时传达信息 [266]。为实现这些目标，专业人员开发了多种监测系统，包括基于病例、事件、实验室、症候群和网络搜索的监测方法 [267]。这些系统依托多种数据来源，例如死亡率和发病统计数据、临床记录、实验室报告、现场调查、人口统计信息和环境数据，建立分析模型，以支持实时监测和态势评估 [268]。为从这些数据中得出准确的监测结果和可付诸行动的洞察力，多个学科的研究人员作出了重要贡献 [269; 270; 271；272]。例如，一些物理学家利用现有数据估算系统参数并计算关键指标，从而构建能够捕捉流行病传播机制的动力学模型 [273]。这类动力学模型为量化传播过程、预测疫情发展轨迹，以及识别高风险人群或区域提供了理论基础，进而提升传染病检测和监测的精准度与时效性。

随着互联网和智能设备的快速发展，大量多源监测数据已唾手可得，例如流行病学报告、出行聚合数据、匿名化数字轨迹，以及与流行病相关的社交网络数据和网络搜索数据 [274]。这些异构数据源有助于更准确地进行流行病检测与监测；然而，传统方法（如动力学建模）难以有效整合和利用此类复杂数据，导致及时性和灵敏度受限 [273]。相比之下，人工智能驱动的方法，尤其是由大语言模型增强的方法，在融合、解读和推理这些多源流行病数据方面具有巨大潜力。在本小节余下部分，我们将综述流行病不同阶段的生物学疫情检测与监测进展。我们首先讨论早期检测 [275]，随后介绍持续疫情监测。

4.2.1. 早期检测

流行病传播的早期检测旨在尽可能早地识别疫情暴发的起始阶段，理想情况下是在广泛传播发生之前，从而为疫情预警奠定基础 [276；277；278； 279]。此类检测对于有效的公共卫生响应和疫情遏制至关重要，有助于及时发布警报并采取管理措施，减少传播并促进早期干预 [280]。接下来，我们按照方法范式对现有研究进行分类。我们首先简要概述用于早期检测的传统基于模型的方法，以及基于机器学习或深度学习的方法。随后，我们将详细综述由大语言模型增强的早期阶段生物学疾病检测方法。

（a）传统的基于模型的检测方法。

这类方法大致可分为时间预警模型、空间预警模型、时空预警模型和动态模型。时间、空间和时空模型旨在通过分析时间趋势、空间分布或其随时间和空间的联合演变，检测潜在的疫情暴发热点。相比之下，动态模型侧重于捕捉人群中的传播机制，以推断潜在的流行病动态。这些方法通常依赖与疾病当前负担和流行病学、病例的时间变化以及疫情暴发和新兴病原体识别相关的数据 [273]。关于此类传统的基于模型的检测方法，更多详情见参考文献 [281；282；283；284； 285]。

（b）基于机器学习或深度学习的方法。

随着互联网和智能设备的广泛普及，新闻文章、社交网络帖子、在线论坛讨论以及公共卫生公报或报告等开放源数据，已成为传统预警系统不可或缺的补充。这些数据可能包含微弱但及时的疫情暴发信号，从而支持近实时态势感知。利用机器学习、深度学习和自然语言处理，人们开发了一系列人工智能驱动的预警系统，从此类异质来源中检测早期流行病信号 [286； 287]。代表性示例包括开放源疫情情报（EIOS） [288； 289], EPIWATCH [290], 以及波士顿大学新发传染病中心（BEACON）系统 [291；292； 119]。

其中，世界卫生组织（WHO）开发的 EIOS 运用机器学习和多语言文本挖掘技术，从新闻文章、社交网络帖子、官方公告和在线健康论坛等多种非传统来源中识别潜在的疫情暴发信号。EIOS 通过自动对开放源报告进行分类和聚类，支持传染病事件的早期检测，并将可能的信号转交专家核实。该系统已部署于 160 多个国家，通过实时监测和多语言适应能力提升全球疫情情报。

EPIWATCH由新南威尔士大学开发，利用机器学习和自然语言处理，持续扫描全球在线数据（如新闻、博客和社交网络），以发现异常疾病模式。它会评估检测到的信号的严重程度和可信度，并发布往往早于官方通报的警报。与EIOS相比，EPIWATCH以研究为导向，尤其适用于学术和公共卫生应用，提供了一个可扩展且经济高效的框架，用于早期流行病情报。

[2]https://www.who.int/emergencies/监测

BEACON利用人工智能驱动的事件监测，在人类、动物和环境健康领域检测并分析生物威胁。该系统汇总开源信息，如新闻源、公告和在线健康内容，并将筛选后的信号转交领域专家进行验证。BEACON强调透明度、实时可访问性和专家验证，以提升流行病警报的精准度和及时性。其他具有代表性的人工智能预警系统包括ProMED-mail [293], EPIWATCH [294], 和HealthMap [295]。此外，除这些通用预警系统外，一些研究还直接应用机器学习或深度学习技术，对特定疾病进行早期检测，相关研究见参考文献 [296；297；298； 299]。

（c）大语言模型增强方法

尽管这些人工智能驱动的系统显著提升了早期检测的自动化程度和响应速度，但大语言模型增强方法能够改善语境推理、跨语言理解以及异构数据整合。在近期一项关于眼部感染的信息监测研究中，Deiner等人 [115] 使用从 X、在线论坛和 YouTube 收集的数千条合成及真实世界结膜炎相关帖子，对七种大语言模型进行了评估。他们设计了提示词，要求模型推断与疫情暴发相关的属性，而非仅仅做出“暴发或未暴发”的二元决策；随后，他们通过相关性分析、敏感度/特异度和评定者间信度，将模型输出与专家评估进行了比较。结果显示，GPT-4 和 Mixtral-822B 的性能最为可靠：GPT-4 估计的疫情暴发概率与专家标注的相关性为 0.73，Mixtral 估计的疫情暴发规模与真实病例数的相关性为 0.82。在真实帖子上验证感染病因时，GPT-4 达到了较高的特异性（0.83˜1.00）和中等的敏感度（0.32˜0.71）。这些发现表明，大语言模型能够提取多个流行病学维度，例如疫情暴发概率、疫情暴发程度和可能的病因，而不仅仅是生成粗略的二元预警。因此，大语言模型可作为有效的早期疾病检测工具，在传统报告系统检测到疫情暴发之前，支持态势感知。

Deiner等人 [116] 还探索了利用大语言模型监测推文帖子并评估流行病暴发可能性的做法。他们收集了 12,194 条与某种特定传染病相关的推文，并将其提供给两个大语言模型（GPT-3.5 和 GPT-4），提示模型评估是否可能发生疫情暴发。结果表明，大语言模型能够有效解读社交网络帖子，达到与人工评估相当的预测准确率。尽管其性能与人工评估相近，但这种方法所需的人力显著更少，因为只需向模型提供相关的社交网络帖子。这表明，大语言模型有望成为早期流行病预警工作流程中一种可扩展、低成本的组成部分。

此外，Kaur 和 Butt [117] 提出了一种用于疫情智能的 LLM驱动的集成框架，该框架整合了多源数据流（包括新闻报道、社交网络内容和网络搜索日志），并配备了资源优化和跨语言处理模块。该框架着重于实时分析、跨源数据融合以及态势评估的自动生成，使其能够在早期阶段检测弱疫情信号和新兴威胁。他们的研究系统阐述了大语言模型在提升早期检测、态势感知和响应准备方面的潜力，同时也指出了若干关键挑战，例如系统架构设计、模型验证、跨语言适配和误报控制。

4.2.2. 持续疫情监测

对于研究生物疾病传播的研究人员而言，高效分析和处理相关数据至关重要，这有助于实现持续监测，并生成能够为疫情趋势评估和有效干预策略的制定提供切实参考的洞察力。多种异构数据也可用于持续疫情监测，既包括医院急诊科收治人数和非处方药销售量等传统公共卫生指标，也包括网络搜索查询和社交网络活动等现代数字踪迹 [300]。与早期检测一样，我们首先简要概述传统的基于模型的方法，以及用于持续疾病监测的机器学习或深度学习方法。随后，我们将详细综述由大语言模型增强的持续疫情监测方法。

(a) 传统的基于模型的方法。

持续疫情监测可借助生物疫情建模实现。例如，Han等 [301]模拟了低收入和中等收入国家中典型的COVID-19疫情场景，以考察检测率、采样策略和测序覆盖率如何共同影响监测结果。研究结果表明，较低的检测率和时空偏差可能使新变异株的识别延迟数周至数月，并可能导致变异株流行率估算不可靠。Allard [302]介绍了自回归积分滑动平均模型在传染病监测中的实际应用，探讨了从模型设定与转换到参数估计、预测和迭代更新的完整工作流程。更多类似研究见参考文献 [303; 304; 305]。

(b) 基于机器学习或深度学习的方法。

随后，许多研究人员开始运用人工智能支持持续监测。这些方法整合嘈杂且复杂的数据，并将其转化为有效信息，从而从难以解读的多维数据中提炼出有意义的洞察。这些人工智能驱动的方法有助于追踪发病率、死亡率、传播来源及其他关键流行病学指标，进而推动传染病监测的发展 [244]。例如，在新冠疫情期间，研究人员通过分析移动电话移动性数据、社会

网络内容和搜索引擎查询日志，从而识别新冠病毒传播信号[3] [266]。Li 等人 [306] 通过结合报告病例观测数据与人类流动数据，对 SARS-CoV-2 进行持续疫情监测。他们构建了网络化动态元种群模型，并采用贝叶斯推断来推断关键流行病学特征，包括未报告感染的比例及其相对传染性。Wu 等人 [307] 在利用知识图谱探索传染病与患者之间关系时，考虑了可能产生的健康信息泄露风险 [308； 309]以开展持续疫情监测。他们提出了一种基于联邦学习 [310 的框架 311]，构建局部差分隐私 [312 的个体症状关系图 313]，用于疫情风险监测；该方法在保护隐私的同时，有效地表示个体之间潜在的传播关系。

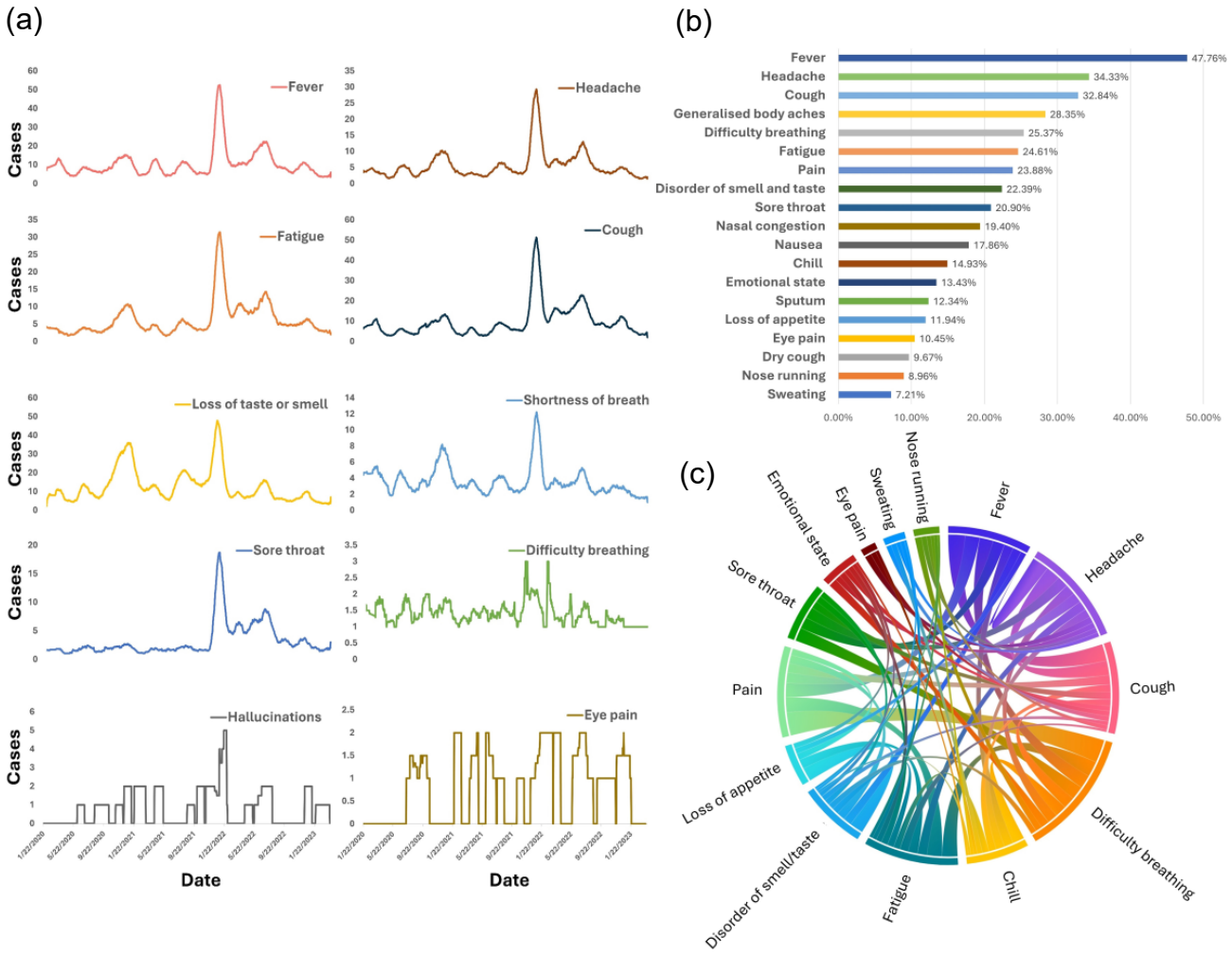


图11：利用大语言模型进行持续疾病监测的示例。大语言模型用于识别社交媒体帖子是否对应感染检测呈阳性的用户，并提取其报告的症状，随后汇总这些信息进行统计分析。(a) 症状流行率随时间变化的趋势。(b) 报告病例中不同症状所占比例。(c) 症状之间的相关性。来源：根据参考文献 [118] 重现

(c) 大语言模型增强方法。

大语言模型在传染病监测方面具有巨大潜力 [314]，，因为它们能够准确处理和分析大量非结构化文本，并帮助专业人员简化劳动密集型工作流程、发现潜在趋势 [244]。Xie等人 [118]观察到，许多出现轻微新冠症状的个体会在家中进行自我检测，并在社交媒体上报告自身状况。为有效利用这些数据，他们使用大语言模型识别和分析与COVID-19感染相关的帖子，从而持续监测疫情发展和症状演变。具体而言，他们首先从平台X收集可能相关的帖子，然后手动标注其中一部分能够证实感染的帖子，并利用这些带标签的示例对大语言模型进行微调。随后，使用微调后的模型检测可能的阳性病例，并提取报告的症状，将这些信息汇总后进行统计分析和实时监测。如图11所示，这一大语言模型增强的监测框架能够实现：（i）追踪症状的时间趋势，从而揭示新变异株的出现（例如，味觉或嗅觉丧失在2020年9月前后变得普遍，这与贝塔变异株的兴起相吻合）。（ii）估计症状流行率，为资源分配和

[3] 疾病传播、疫苗接种率及相关过程的分析通常依赖汇总数据、匿名数据或其他保护隐私的数据，并在既定的公共卫生和伦理框架内开展。

准备工作规划提供支持。（iii）分析症状之间的相关性，为临床进展模式提供洞察力。此外，他们的纵向研究结果表明：a）从2020年到2023年，康复时间逐渐缩短；b）2023年的大多数COVID-19病例症状较轻，表明病毒严重程度有所下降；c）180天内再感染的情况相对不常见。d）不到10%的个体经历过两次再感染，约1%的个体经历过三次感染，而报告四次或更多次再感染的比例则更低。Consoli等人 [119]系统评估了九种大语言模型在持续疫情监测任务中的能力，考察它们从非结构化文本中提取疫情相关信息的性能。他们进一步应用上下文学习来增强模型能力，使准确性和处理效率均得到显著提升。Sufi等人 [120]探索了使用大语言模型进行新冠监测。他们同样利用大语言模型从社交网络中提取和分析与疫情相关的信号。

5. 流行病预测与管 理：大语言模型增强的传 播动力学 s

研究传播动力学的最终目标之一，是实现准确预测和有效干预，从而缓解或防止不良传播。 [149; 5]. 本节对大语言模型在传播过程预测与管理中的应用研究进行分类和综述。具体而言，本节回顾大语言模型增强的数字流行病和生物疫情研究，概述大语言模型如何助力预测和管理传播。每个小节首先考察大语言模型如何促进流行病预测，然后讨论其如何推动流行病管理。

5.1. 大语言模型增强的数字流行病监管

在物理学，尤其是统计物理学、复杂系统和非线性动力学领域，信息传播并非像传播学那样从媒体和受众的视角加以研究，而是被视为复杂网络上的一种动力学过程。研究重点在于信息、谣言或观点如何沿着网络的“节点—边”结构扩散。这一过程通常被视为一种渗流样过程，并与理论物理学中的相变相类比 [315]。常见的建模方法包括 SIR/SEIR流行病模型、阈值模型 [316], 谣言模型（例如 Daley-Kendall模型、Maki-Thompson模型）、投票者模型、伊辛模型和自旋动力学、渗流理论、独立级联模型 [317], 以及基于主体的仿真。研究主题涵盖不同网络拓扑（小世界、无标度、社群结构）中的传播规律；识别有影响力节点以最大化传播范围；抑制错误信息或病毒扩散的策略；同一网络中多条信息或观点之间的竞争与共存；网络结构的时间演化；以及影响传播效率和稳定性的因素。物理学家将信息扩散视为一种动力学过程，并通过建模、分析和模拟，力求理解、预测并最终引导传播，或在更广泛的语境下对其进行管理。本小节首先简要综述信息传播预测的传统方法，随后详细讨论基于大语言模型的预测方法，以展现从经典模型到大语言模型驱动范式的演进。

在大语言模型出现之前，传统的信息传播预测方法主要依赖 SIR 和 SEIR 等数学模型。这些模型不依赖特定数据集，能够应用于任意复杂的网络，因此具有很强的泛化能力。然而，这些模型依赖较强的假设，难以捕捉现实社交网络环境中的异质性。它们可使用的输入信息有限，通常只关注网络结构和初始感染节点，处理文本内容和语义信息的能力也较为有限。例如，Cao 等人 [318] 利用早期采用者的状态以及底层网络结构，预测其他节点是否会被感染。具体而言，他们同时考虑目标用户及其邻居的激活状态，以及各自的影响力，对节点的激活状态进行建模和预测。如图 12 所示，在已知节点 A 和 E 已感染且其在网络中的位置已知的情况下，他们采用图神经网络迭代预测哪些节点可能会成为下一批感染者。

Wu 等人 [319]研究了信息扩散宏观演化预测问题。他们将信息扩散归因于社会影响力，认为社会影响力驱动信息从一个节点转移到另一个节点。基于这一洞察力，他们提出了一种定向标签传播算法，将个体节点的微观传播模块进行聚合，以预测宏观扩散模式，如图 13 所示。在其框架中，节点是否成为感染者被视为一个标签，而标签选择机制对应于传播的微观决策过程。通过对微观行为进行建模，他们识别出了潜在的影响因素和扩散的关键驱动机制。有趣的是，他们发现，增加特征数量并不一定能提高扩散行为估计的准确性。此外，当考虑了足够数量的特征时，估计准确性与预测模型的选择并无显著相关性。另外，Taylor 等人 [320]和 Jin 等人 [321], 注意到，社群内的信息传播不同于个体层面的信息传播，因此通过处理节点群组之间的预测，将研究视角从个体层面扩展到了社区层面。尽管他们的方法纳入了文本信息，但处理此类内容的能力仍然有限。

与传统预测方法相比，GPT、LLaMA 等大语言模型的出现改变了信息传播预测的方法。大语言模型由数据驱动，能够处理海量社交网络内容，并从文本、图像和视频中提取主题与情感。它们不仅能够预测数字疫情是否会爆发，还能预判哪些主题会传播、在什么群体中传播，以及传播时所带有的情感立场。通过整合文本语义与网络结构，大语言模型能够处理讽刺和隐喻等复杂语境，

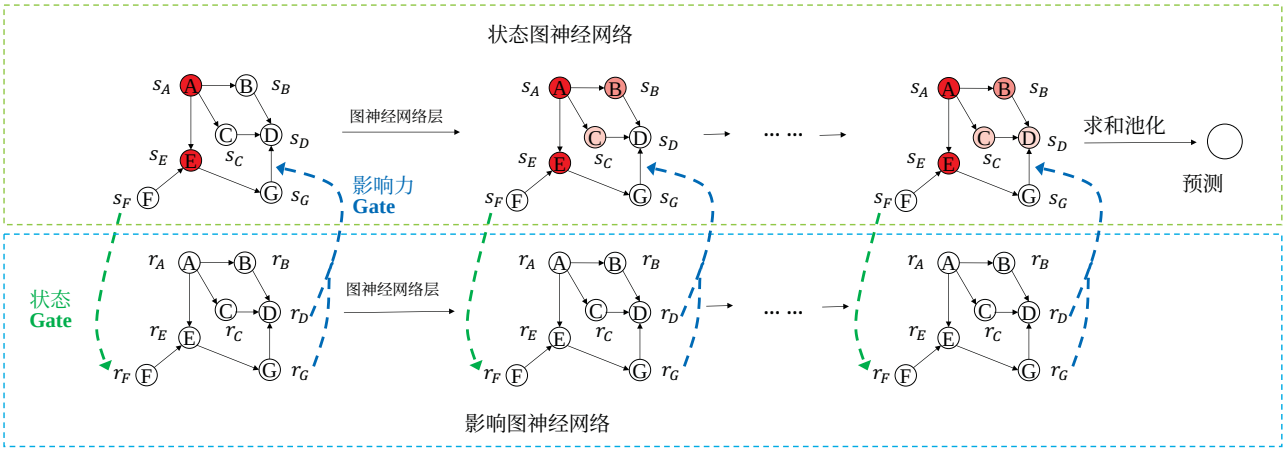


图12：示例：利用早期采用者的状态及底层网络结构，预测其他节点的感染情况。在由七个节点构成的有向网络中，每个节点都关联一个初始感染状态，而有向边则编码节点之间的影响关系。该方法采用图神经网络，根据底层网络结构、初始节点状态和有向交互来学习传播动力学。基于学到的传播模式，模型预测其他节点在后续时间步的感染状态。来源：根据参考文献 [318] 转载

从而实现更准确的预测。大多数信息传播预测与干预方法大致可分为两类：使用大语言模型直接预测（不涉及建模过程），以及基于第 3 节所述建模过程的基于模拟的预测。

5.1.1. 用于预测的大语言模型

(a) **利用大语言模型进行预测**

如前所述，与传统方法相比，基于大语言模型的方法在处理文本内容和语义信息方面能力更强。大语言模型可以被视为一个黑箱：一旦将文本信息作为提示输入，它便会展现出语义理解能力，并反映在其输出中。在信息传播预测这一背景下，一个关键问题是如何利用大语言模型捕捉此类语义信息。为此，我们以 Wang等人的研究 [121]为例，该研究探索了利用大语言模型从新闻文章中提取内容以进行时间序列预测。如图14所示，研究人员通过文本提示将新闻及补充信息（如天气、日历日期和地理数据）嵌入时间序列数据中，随后微调预训练大语言模型，以预测序列中的数值序列。微调过程首先检索相关的原始新闻和补充信息。然后，基于大语言模型的推理智能体会迭代筛除与预测任务无关的新闻，并从相关事件中提取有价值的洞察力。接下来，将筛选后的新闻与提示模板结合，形成提示词，作为微调后大语言模型的输入。随后，模型预测时间序列中的数值序列。最后，该框架会评估预测结果的准确性，并通过迭代不断优化新闻筛选逻辑和智能体输出的稳健性。大语言模型完成微调后，后续预测只需将相应的新闻和时间序列任务输入模型，即可自动生成预测结果。设历史数据集为 $\mathcal{H}$，历史提示集为 $\mathcal{Q}$，历史输出集为 $\mathcal{O}$。将训练数据集定义为 $\mathcal{D} = \{(h_i, q_i, o_i)\}^{|\mathcal{H}|}_{i=1}$，其中$h_i \in \mathcal{H}$ 表示历史数据，$q_i \in \mathcal{Q}$ 表示历史提示词，$o_i \in \mathcal{O}$ 表示相应的输出。给定参数为 θ的预训练大语言模型 $\mathcal{M}_\theta$，微调过程可表述为最小化训练目标

$$\begin{aligned} \hat{\theta} &= \arg\min_{\theta'} \mathcal{L}(\mathcal{M}_{\theta'}; \mathcal{D}), \\ \hat{\mathcal{M}} &= \mathcal{M}_{\hat{\theta}}, \end{aligned} \tag{12}$$

其中，$\mathcal{L}$ 是一个合适的损失函数，用于衡量预测输出与目标输出之间的差异。微调后，对于新的预测任务，会将提示q 作为输入提供给微调后的模型 $\mathcal{M}$，然后由该模型通过$\hat{o}^q = \mathcal{M}(q)$生成预测。

接下来，我们将详细概述基于大语言模型的信息传播预测方法。在这一任务中，大语言模型通常会将其重新表述为未来节点激活预测，即给定已感染的节点和底层网络结构，预测接下来哪些节点会被感染。Zhong 等人 [122]提出了一种名为级联检索上下文学习（CARE）的框架，该框架利用大语言模型推断社交网络中用户接下来是否会被激活。在这里，级联指信息从一个用户传递到另一个用户所经过的路径。CARE 包含三个关键组件：提示池构建、提示检索，以及用于预测的上下文学习。首先，CARE 从历史级联中提取并构建提示池，其中包含各种信息传播模式。这些提示作为传播行为的过往示例。接下来，系统利用基于排序的搜索引擎技术，从提示池中检索与目标级联传播模式相似的提示，从而找出扩散机制与当前传播过程相似的历史案例。最后，将检索到的提示整合为上下文输入

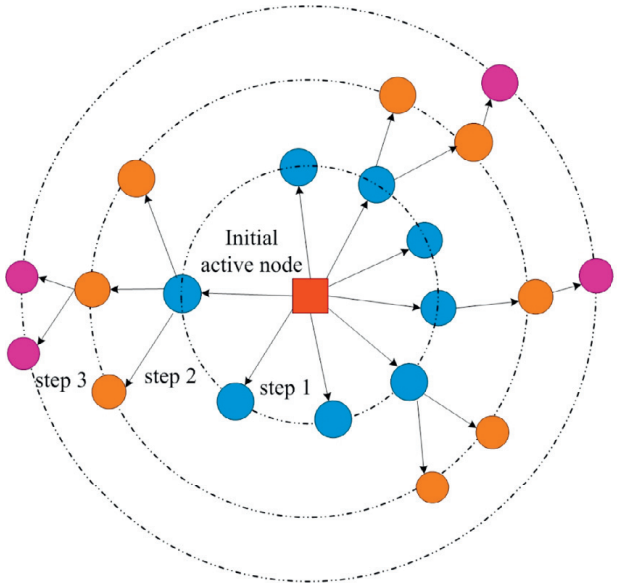


图13：定向标签传播示例。该网络由多个节点组成，其中只有位于中心的矩形节点最初处于激活状态，其余圆形节点均处于不活跃状态。根据不同颜色，将不活跃节点划分为多个层。同一层中的节点被视为相互独立，并同步、独立地更新其状态。每一层对应离散时间间隔内的一轮感染。在每一轮中，所有易感节点（定义为活跃节点的不活跃邻居）都会根据微观行为估计模型更新其状态。通过逐层进行此类估计，可以模拟传播过程，从而分析扩散的潜在影响因素和关键驱动机制，并预测后续时间步中可能被感染的节点。来源：转载自参考文献 [319]

传递给大语言模型，由其对目标信息进行传播预测，例如推断个体用户的感染行为。与时间序列预测相比，CARE 中的训练数据集 $\mathcal{D}$还额外包含网络结构 G，即 $\mathcal{D} = \{\mathcal{H}, \mathcal{Q}, \mathcal{O}, G\}$。

Shang 等人 [123]针对如何使大语言模型有效地从网络结构和用户档案文本中提取信息这一挑战，提出了一种融合式注意力感知预测方法。他们使用受约束的广度优先搜索算法从网络结构中提取级联路径，为级联建模提供结构基础。将用户帖子与这些路径整合，构造用于微调 Diffusion-LLM 的训练提示，使其能够捕捉结构信息和语义信息，并识别具有扩散影响力的用户对。用户档案通过 Transformer 编码，以生成相似度矩阵。最后，将 Diffusion-LLM 生成的扩散影响力矩阵与用户档案生成的相似度矩阵融合，再将所得表示输入多头注意力层，生成用于预测的级联表示。

Chen 等人 [124] 针对字段名与字段值之间的语义关系常被忽略、导致上下文信息丢失且特征表示不足的问题，提出了一种将元数据转换与基于 LoRA 的大语言模型微调相结合的方法。元数据转换会将结构化的社交网络元数据（例如 UID、帖子标识符、地理精度、标题）转换为语义丰富且包含上下文信息的自然语言描述。例如，一条“Uid: 25893@N22, Pid: 565381, Title: Beautiful tree.”记录会被转换为“A social network user (5893 @N22) has uploaded a post (565381) with a geo-accuracy of 16, titled ‘Beautiful tree’”。这种转换明确表达隐含含义，并关联不同的元数据元素。随后，使用 LoRA 技术对预训练大语言模型进行微调，以完成预测任务。

(b) **将大语言模型作为组件以增强预测能力**

与直接依赖大语言模型生成预测输出的方法不同，有些方法利用大语言模型辅助训练专用预测模型。Gao 等人 [126]采用这一策略来预测视频在社群中的传播。如图 15 所示，他们的框架由五个模块组成：视频理解、视频增强、社群画像

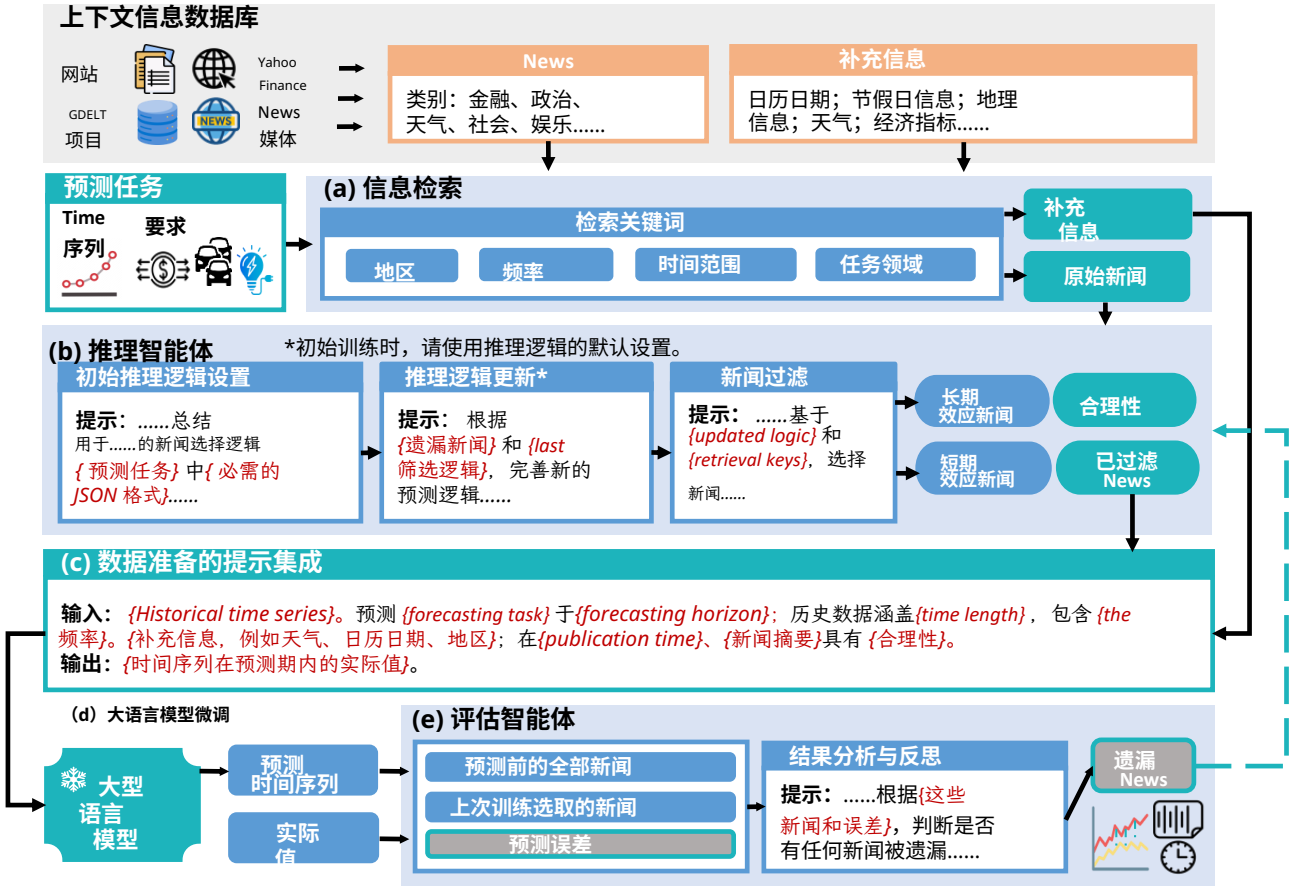


图14：大语言模型在时间序列预测中整合文本信息的示例。通过微调，使基础大语言模型适用于特定的时间序列预测任务。整体工作流程包括以下步骤。(a) 根据预测任务的地理位置、时间范围和上下文要求，从互联网或集成的专有数据库中检索相关的原始新闻文章和补充信息。(b) 基于大语言模型的智能体分析并筛选相关新闻，过滤掉与预测任务无关的新闻。(c) 将筛选出的新闻与预定义的提示模板结合，构建特定任务的提示词。(d) 随后使用这些提示词对大语言模型进行微调，并将其用于执行预测任务。(e) 当模型的预测结果与观测到的真实值存在差异时，将触发对历史新闻和数据的回顾性综述，从而改进推理过程。来源：转载自参考文献 [121]

构建、动态公共链接和预测。在视频理解模块中，多模态大语言模型处理原始视频文件，生成文本描述，再将其与视频标题、字幕和用户提供的描述相结合，形成统一的视频文本表示。该表示被编码为融入语义信息的低维向量。在视频增强模块中，大语言模型将相关的历史、文化和社会知识注入视频表示，为后续预测提供更丰富的语境基础。在社区画像构建模块中，大语言模型分析每个社区，构建特征画像，并将其表示为社区特征向量。在动态公共链接模块中，对活跃度较低的社区进行特殊处理，以提高鲁棒性。最后，在预测模块中，将视频文本表示和社区特征表示作为输入进行整合，以训练轻量级预测模型，预测社区级视频传播。总而言之，大语言模型用于视频理解、视频增强和社区画像构建模块，将视频和社区特征转化为语义丰富的表示，再将这些表示用作最终预测模型的输入。重要的是，大语言模型本身并不充当预测器，而是作为特征提取器和知识整合器，为下游预测提供支持。

同样，Sun等人 [127]提出对预训练大语言模型进行适配，用于转发预测，从而提高社交网络中转发行为建模的准确性。他们的框架通过两个关键步骤发挥大语言模型的优势：多模态特征提取和语义时间适配。第一步，大语言模型从文本、视觉内容和图结构中提取多维特征。为将通用领域知识迁移至转发预测任务，他们在每个大语言模型模块的自注意力层之后插入轻量级适配器，引导模型捕捉表明转发倾向的语义模式。第二步，他们在每个前馈层之后设置时间适配器，以捕捉用户兴趣和偏好的动态变化。用户历史记录在各个时间适配器间进行分段，并通过多头机制分别适配每个片段，随后合并输出。通过结合基于大语言模型的多模态表示以及语义和时间适配，该方法显著提高了预测准确率，并克服了现有方法在处理稀疏数据和高维时间特征方面的局限。

(c) **利用大语言模型生成智能体以模拟扩散**
这种方法　　d 利用大语言模型生成智能体，在模拟环境中复现用户。通过建模

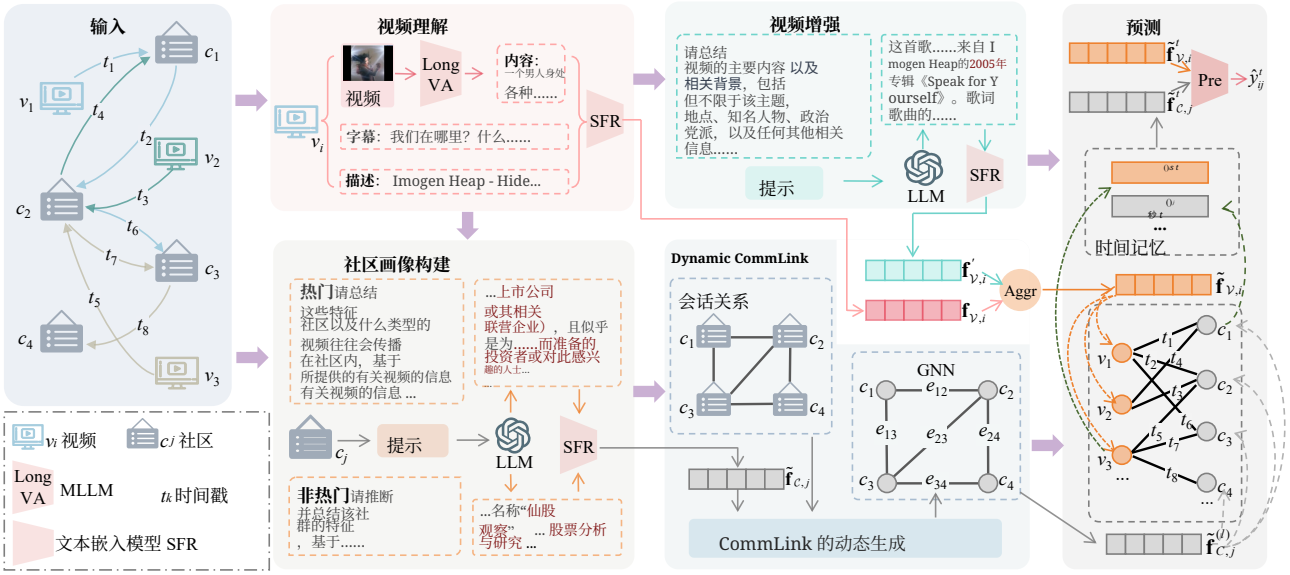


图15：展示将大语言模型作为组件以增强预测能力。该框架由五个模块组成。不同的大语言模型被用作不同的工具：它们分别将视频内容和社区信息编码为文本视频表示和社区特征表示。随后，这些表示向量被用作输入，以训练专门的预测模型。来源：改编自参考文献 [126]

通过这些智能体之间的交互和信息传播，系统能够再现传播动力学，并预测未来的扩散。Gao 等人 [128] 提出了一种由大语言模型智能体增强的社交网络仿真系统，旨在模拟真实社交平台上观察到的传播过程。如图 16 所示，现实世界数据（包括用户、关系和消息）被输入模拟环境，在该环境中，由大语言模型赋能的智能体会生成点赞、转发、评论和发帖等用户级行为。这些行为经过迭代更新，并汇总为信息、态度和情绪扩散的群体级动态，从而实现对社会过程的预测和评估。该系统包含三个主要组件：(i) 环境构建。从真实社交平台收集用户数据、关注者关系和文本帖子，以构建有向图。用户被划分为高影响力、普通和低影响力群体，以反映其在转发和交互行为方面的差异。系统使用带有提示和微调的大语言模型来推断并补全用户属性（例如年龄、性别和职业），且每个智能体都维护一个记忆池，依据时间衰减和内容相关性等因素筛选有影响力的帖子。(ii) 个体级模拟。在每个时间步，智能体根据其用户档案和记忆池所提供的信息，通过马尔可夫过程更新情绪状态（平静、中等、强烈）和态度（积极、消极）。基于这些状态，智能体决定是否转发、生成原创内容（文本由大语言模型生成）或保持不活跃。(iii) 群体级动态。汇总个体行为后，可生成信息、情绪和态度的宏观扩散曲线，并将其与现实世界数据进行比较，以评估模拟的保真度。通过这种设计，系统将环境构建、个体模拟和群体级汇总整合到统一框架中。该系统可用于预测社会运动、公众舆论变化或新文化实践的采纳情况。

类似地，Chuang等人 [66] 采用基于大语言模型的多智能体系统替代传统的基于主体的模型，以模拟人类观点动力学。通过这一模拟，模型预测群体观点的演化轨迹，并考察初始意见分布、网络规模和偏见强度等因素如何影响共识或极化结果。例如，模拟指标可用于预测：偏见被定义为群体观点的平均值，反映共识可能形成的方向；多样性被定义为观点的标准差，表示观点极化的程度。

Li 等人 [62] 采用LLM驱动的多智能体仿真框架，对不同社交网络结构中的虚假新闻扩散动力学进行建模，由此产生的模拟结果能够预测虚假新闻传播趋势。他们研究了三类网络（随机网络、无标度网络和高度聚类网络）的影响，以及基于大五人格特质 [322] 的不同智能体人格。结果显示，这些因素会影响传播动力学：虚假新闻在无标度网络中传播最快，在高度聚类网络中传播最慢；而某些人格特质（如外向性和开放性）会提高新闻分享的可能性。

Ma等人 [68] 提出了一种模拟错误信息易感性测试（SMIST）方法，用LLM生成的对应个体取代人类参与者。SMIST通过构建生活经历和人口学画像来创建模拟参与者，随后向这些智能体发放有关错误信息的问卷。其核心目标是评估个体对错误信息的易感性。通过生成具有不同人口学属性的参与者，例如年龄、性别、教育水平、政治倾向，以及对科学家、政府和新闻业的信任，SMIST能够预测这些因素如何影响个体相信并传播错误信息的倾向。

Yao 等人 [70] 指出，传统算法难以捕捉现实社会互动的复杂性，尤其难以刻画大幅度的观点波动和自衰减现象。为此，他们提出了 FDE-LLM，这是一种混合方法

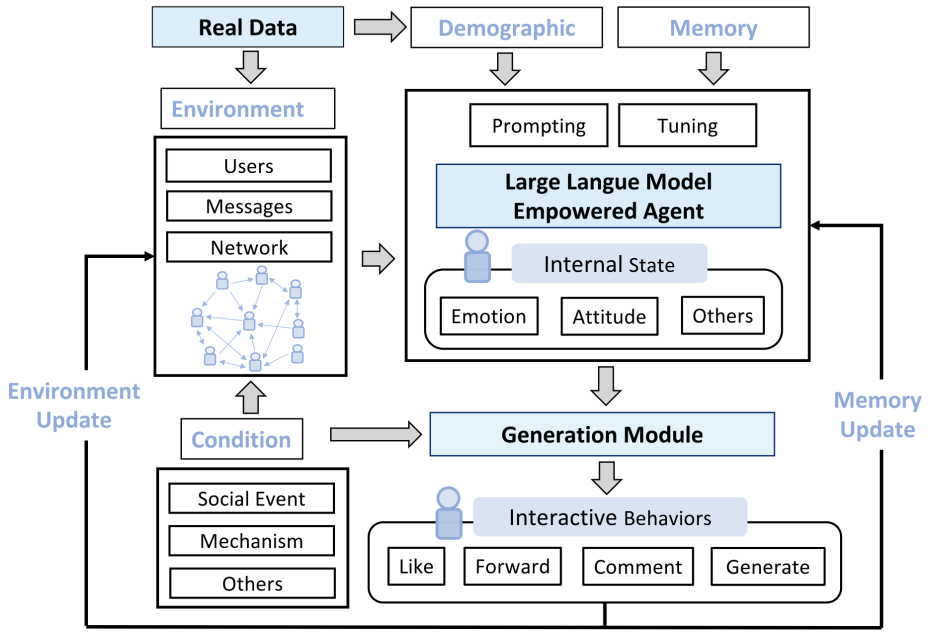


图 16：利用大语言模型生成智能体，以模拟传播动力学并实现预测的示意图。现实世界数据被输入模拟环境，由大语言模型赋能的智能体复现用户行为。这些行为经过迭代更新和汇总，从而支持对信息传播的预测。来源：转载自参考文献 [128]

旨在更准确地模拟和预测社交网络中的观点动力学。在该框架中，用户被分为两类：意见领袖和意见追随者。意见领袖采用大语言模型与元胞自动机（CA）相结合的方式进行建模，而意见追随者则通过 CA 和 SIR 模型进行模拟。大语言模型能够有效捕捉剧烈的观点转变，CA 则考虑了人际影响和连续动态，SIR 模型模拟用户对重复事件的“疲劳”状态，使观点逐渐衰减至中立。实验结果表明，在各项评估指标上，FDE-LLM 的表现均显著优于传统的基于主体的模型和纯大语言模型方法，展现出更高的准确性和鲁棒性，能够更有效地模拟观点动力学。这些发现凸显了将大语言模型与传统动态模型相结合的潜力，为社交平台上的观点预测提供了更精确的解决方案。

Rossetti 等人 [129] 提出了 Y Social，这是一个由大语言模型驱动、面向社交网络的数字孪生平台，可模拟复杂的智能体行为，包括用户互动、内容传播和网络动态。通过对内容传播和结构演化进行细粒度模拟，Y Social 为预测用户回应，进而预测信息扩散奠定了基础。Zhang 等人 [130] 将大语言模型集成到基于主体的模型中，以更准确地模拟社交网络中的影响力扩散。他们的方法包含两个主要组成部分：大语言模型增强型智能体模型和模拟执行。前者扩展了传统的独立级联（IC）模型；在传统模型中，智能体仅有二元状态（活跃或不活跃），且传播以固定概率发生。扩展后的模型融入了大语言模型能力，以捕捉个体智能体的异质性。后者负责管理整体扩散过程并开展数据分析。随后，该模拟框架被用于预测与新闻、观点和主题相关的用户行为。Liu 等人 [131] 提出了 MOSAIC，这是一个基于大语言模型的多智能体社交网络仿真框架，用于预测点赞、分享和举报内容等用户行为。该框架首先利用大语言模型生成具备记忆、反思和推理能力的智能体。通过预先设定的问答提示和场景引导，为每个智能体赋予人格特质和行为动机。随后构建有向图来表示真实社交平台（例如 X），其中节点代表用户，边代表关注关系，从而能够模拟内容扩散路径。通过互动，智能体基于动态记忆形成行为模式，复现多样化的人类行为。最后，MOSAIC 纳入三种错误信息管理机制，以便在不同内容审核策略下调整智能体行为，从而分析内容扩散动态、传播速度和范围，以及各项策略在抑制错误信息的同时维持用户参与度的有效性。

5.1.2. 用于管理的大语言模型

除了有助于理解信息扩散的原理、检测和监测新兴事件，以及预测其传播路径、持续时间和轨迹之外，关于信息传播的研究在管理中也发挥着关键作用，通常体现为管理或干预。这对于有害内容的监管尤为重要，其中包括恐怖主义宣传、极端主义材料、错误信息和有害谣言等。例如，为防止恐怖主义宣传、极端主义内容和仇恨言论在网上不受控制地传播，欧盟委员会要求托管服务提供商在收到删除令后立即删除恐怖主义内容，且无论如何不得超过一小时 [323]。同样，世界卫生组织正式提出了信息疫情管理这一概念，指系统地应用基于证据和风险的干预措施，以预防、检测、

并应对卫生突发事件期间错误信息和虚假信息造成的有害影响 [324]。管理可从互补的视角加以考察，尤其是监管视角和干预策略视角。从监管视角来看，管理包括检测有害内容的传播并向相关部门报告，由相关部门指示托管服务提供商删除相关内容。从干预策略视角来看，平台和监管机构关注的是如何设计更有效的干预措施。一个核心研究问题是，知名学者、专家或主流媒体机构开展的事实核查工作在什么条件下最能有效缓解或消除谣言。在这方面，越来越多的研究探讨了如何利用大语言模型，更好地理解并优化干预策略。

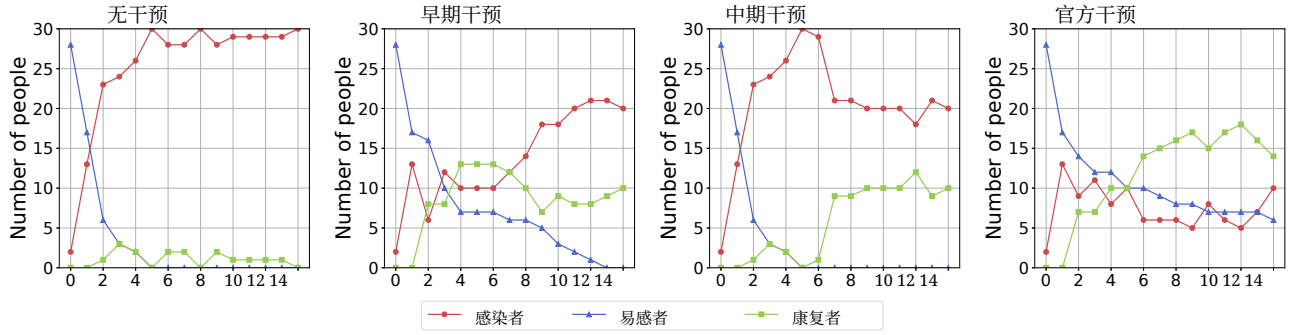


图17：不同干预策略的比较。在模拟的假新闻传播环境中，对30名用户进行建模，并追踪一条政治新闻条目在14天内的扩散情况。从左至右：第一个子图展示了未采取干预措施的基线场景；第二个子图展示了从第1天开始的早期干预；第三个子图展示了从第7天开始的中期干预；第四个子图比较了官方干预措施每日或每三天发布一次事实核查新闻的效果。来源：转载自参考文献 [63]

Liu 等人 [63] 提出了一种由基于大语言模型的智能体驱动的假新闻 FPS，用于模拟个体的态度动态和假新闻传播。每个智能体都具备独特的人格特质、记忆和反思机制，以近似模拟人类认知与行为。在这一框架下，他们考察了政治议题中的官方干预策略，重点关注干预的两个维度：时间和频率。如图 17 所示，在干预时间方面，实验比较了早期干预（在疫情暴发第 1 天实施）与中期干预（在第 7 天实施）。早期干预显著减少了最初的相信者人数，但由于遗忘，其效果会随时间推移而减弱。相比之下，中期干预的抑制效果更为持久，但由于许多用户已受到影响，仍有相当一部分人群受到波及。在干预频率方面，研究评估了官方事实核查新闻每日发布与每三天发布一次的效果，发现两者差异很小。这些发现表明，以适中间隔开展事实核查，是遏制假新闻传播的一种有效且经济的方式。

Lu 等人[78] 探讨了由大语言模型驱动的对抗性社会影响如何影响人们的在线信息处理能力。基于大语言模型支持的实验，他们提出了应对大语言模型驱动的错误信息的干预策略的未来研究方向。例如，他们建议使用机器人标识、大语言模型能力描述（如流畅的语言表达和广博的世界知识），或标明内容是否由大语言模型生成的标签，以提高用户对机器生成内容的认知和警惕性。

Li 等人 [62] 设计并评估了三种旨在遏制假新闻传播的具体干预策略。第一种是鼓励评论，即提示智能体在分享新闻时发表评论，从而激发批判性思维并减少冲动传播。第二种是披露新闻准确性，即模拟事实核查机制：当新闻扩散到一定程度后，通知智能体该新闻不准确，以防止其进一步传播。第三种是屏蔽极具影响力的智能体，即识别并移除网络中的“影响者”——这些主体具有较高的连通性，以及开放性或外向性等人格特质——从而直接扰乱信息扩散的关键节点。

Almaliki 等人[134]结合大语言模型和说服技巧，检测并干预网络仇恨言论。他们采用了与大语言模型集成的基于模板的提示策略，针对仇恨言论检测和干预建议推荐对模型进行定制。当用户在互动新闻网站上发布评论时，系统会在提交前拦截该评-论，通过将输入与两个预定义模板相结合，执行基于模板的提示工程流程，随后生成针对具体语境的建议，供用户进行干预。

事实上，干预措施大致可分为三种不同类型：助推、增强和反驳策略 [227]。现有采用大语言模型分析干预策略的研究，其范围仍相当有限，且缺乏对各类别的细致考察；与传统研究中更为详尽的探索相比，这一不足尤为明显。以往研究表明，即便在个体层面，干预措施的形式也十分多样，从纠正错误信息 [325] 、提升媒体素养 [326] ，到塑造强调准确性的传播环境 [327]，均涵盖其中。因此，如何更有效地利用大语言模型来支持并推进信息管理策略，仍有很大的研究潜力。

5.2. 大语言模型增强的生物疫情调控

在本小节中，我们首先简要概述非大语言模型场景下关于生物学流行病预测（也称生物疫情预报 [328], ）和管理 [329] 的传统研究，随后详细讨论大语言模型带来的最新进展。

生物疫情预测旨在预判传染病可能的传播范围、规模和动态变化，最终目标是支持有效管理。例如，在新冠疫情暴发期间，卡内基梅隆大学的一个研究团队利用谷歌、Facebook 等来源的数据，对美国全国疫情进行实时预测 [330]。该系统可提前最多四周预测重症监护病房对病床和呼吸机的需求，从而能够主动筹备和调配关键医疗资源。Kissler 等人 [331]利用时间序列数据估算两种人类冠状病毒的季节性、免疫力和交叉免疫，并据此构建SARS-CoV-2传播模型，以开展流行病预测。他们的结果表明，首次大规模暴发后，SARS-CoV-2 很可能会在冬季反复激增。随后，他们提出了两项流行病管理措施：（i）长期或间歇性地让公众保持社交距离；（ii）扩大重症监护能力并提高临床治疗效果。

生物流行病管理涉及设计和实施缓解与引导措施，长期以来一直是世界卫生组织、国家及地区卫生主管部门等机构关注的核心议题。在此背景下，教育、学术和医学社群建立了传染病流行病学这一学科，其主要关注点之一是制定并评估预防、管理和缓解传染病传播的策略 [332]。Kissler 等人 [160] 利用大语言模型强大的文本分析能力，对文本、病例报告及相关数据进行分析。因此，疾病建模与分析可以基于更全面的数据集，并借助先进的分析能力。这类人工智能驱动的方法有望重塑传染病响应措施，并推动制定更具针对性、公平性和有效性的干预措施。

管理措施（如政策或干预策略）通常可分为两类。第一类是由人类设计的干预措施，即权威政府机构或领域专家根据建模与预测结果制定策略。第二类则依赖模型或算法驱动的建议，由计算系统生成干预策略。前者主要强调建模与预测的准确性，相关内容已在前文综述。对于后者，如何有效利用大语言模型生成有针对性且高效的干预策略，仍是一个开放且前景广阔的研究领域。下文将聚焦近期利用大语言模型生成或支持干预建议的研究。

传统上，疫情预测主要依赖两种方法：一是机制模型，例如 SIR 或 SEIR 模型，通过舱室结构描述传播动力学；二是统计模型，从历史数据中学习模式 [83； 333]。例如，Tizzoni 等人 [334] 研究了对不同国家感染人数达到峰值的时间进行预测，这一时间通常定义为报告病例数最高的一周 [335]。他们使用全球疫情与流动模型来预测疫情峰值出现的时间，并证明了准确预测是可行的。如图 18 所示，预测峰值时间与实际峰值时间之间存在很强的相关性：约半数国家的预测误差在 2 周以内，超过 95% 的国家误差在 4 周以内。尽管这些方法有一定用处，但往往无法捕捉现实世界中的关键因素，包括受影响人群异质的人口学画像 [336], 、病毒的病毒学特征 [337], 、推动病毒进化的潜在病毒学动态 [338], ，以及公共政策与人类行为之间不断变化的互动 [339], 等 [83]。这些局限降低了对疾病传播模式进行预测的准确性和可靠性 [340]。

为克服这些问题，近期研究开始利用大语言模型整合多样且复杂的数据来源，包括流行病学报告、生物医学文献，以及行为和政策相关信息，从而改进传染病传播的建模与预测。Du等人 [83] 推出了 PandemicLLM，这是一种经过微调的多模态大语言模型，旨在通过学习异构关联数据来预测疾病的短期传播。这些数据包括：(i) 空间数据，例如人口统计信息、医疗系统评分和政治归属；(ii) 流行病学时间序列数据，包括报告病例、住院情况和疫苗接种率；(iii) 公共卫生政策数据，包括文本政策文件；以及 (iv) 基因组监测数据，包括有关变异株病毒学特征及其流行率的权威报告。传统流行病模型主要依赖结构化数值输入，无法充分利用这类异构信息。PandemicLLM 将流行病预测重新表述为基于文本的推理问题，把所有数据转换成文本形式，使大语言模型能够处理并推理传统方法无法获取的模态。如图 19 所示，PandemicLLM 通过两条互补路径整合多模态输入：任务、空间和时间信息通过提示词转换为文本表征向量，而序列数据则由循环神经网络单独编码为序列表征向量。随后，这些表示被拼接起来并输入大语言模型进行预测。与传统的 CDC集成模型相比，PandemicLLM 将预测准确率提高了 119.7%，并将均方误差（MSE）降低了 72.4% [83]。

同样，Moon等人 [82]采用大语言模型，从文本与时间序列数据中联合学习，以预测流感样疾病（流感样疾病）。他们总共使用了两个大语言模型。第一个用于处理文本数据，其中一个大语言模型将文本数据转换为表示向量。第二个用于预测阶段，另一个大语言模型以表示向量作为输入进行微调，从而获得预测模型。该预测模型的输入还包括一个时间序列表示向量，该向量与文本表示向量进行拼接后，再作为输入传递给用于预测的大语言模型。如图20所示，该模型用于评估提前两周和二十周预测流感样疾病发生情况的能力。

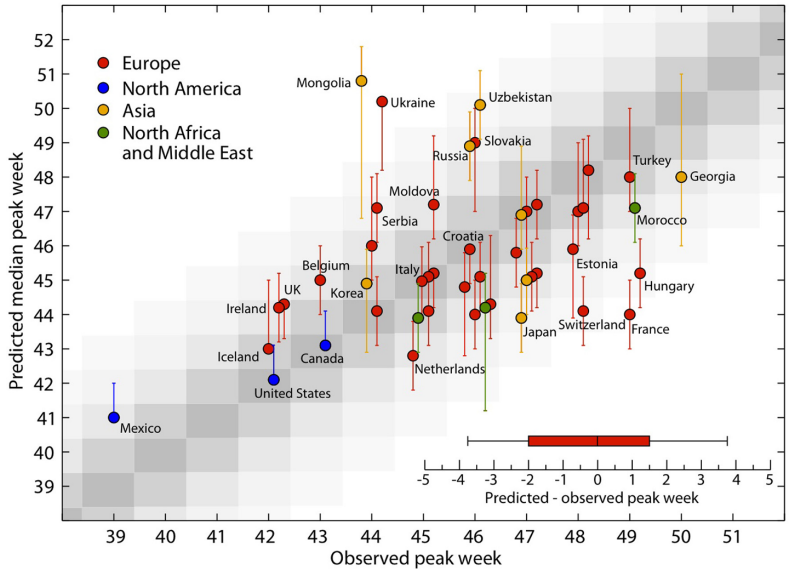


图18：预测与实际疫情高峰期之间的关系。利用2009年H1N1流感大流行期间收集的感染数据，验证了高峰时间预测模型的有效性。x轴表示实际疫情高峰时间，y轴表示预测高峰时间。在欧洲、北美、亚洲以及北非和中东各国，预测高峰时间与实际高峰时间之间可观察到强相关。约一半国家的预测误差在两周以内，超过95%的国家预测误差在四周以内。来源：转载自参考文献 [334]

结果表明，短期（两周）预测具有较高的准确性，可为专家制定疫情干预决策提供有价值的决策支持；而长期（二十周）预测仍有进一步提升的空间。

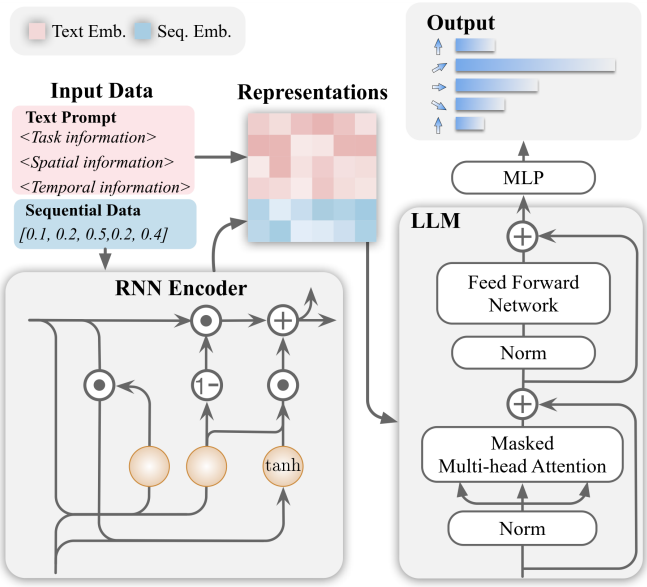


图19：PandemicLLM 中多模态数据的处理。空间、时间和任务相关信息通过提示词转换为文本表征向量，而序列数据则通过 RNN 编码器编码为序列表征向量。随后，将这两类向量拼接并输入大语言模型进行预测。来源：转载自参考文献 [83] 的补充信息

对于仅依赖时间序列数据的场景，Saeed 等人 [132] 提出了 LLM4Cast，这是一个利用预训练大语言模型预测病毒性疾病趋势的框架。该框架对多变量时间序列进行预处理：将其转换为相互独立的单通道序列，以减少通道间噪声；应用 z分数和实例归一化来应对非平稳性；并将数据划分为分块，再通过线性嵌入生成高维表征。该模型整合了双向Transformer编码器和预训练 TinyLlama 模块，训练分为两个阶段：（i）泛化训练：利用跨多个领域的多样化数据集，使 TinyLlama 从文本建模适配为时间序列建模；（ii）领域特定训练：在病毒性疾病数据上微调 TinyLlama。这种分阶段策略构建了一个轻量级大语言模型框架，既能准确预测病毒性疾病，也支持跨领域泛化和高效适配。

上述内容　　这些研究共同表明，利用大语言模型强大的分析能力进行

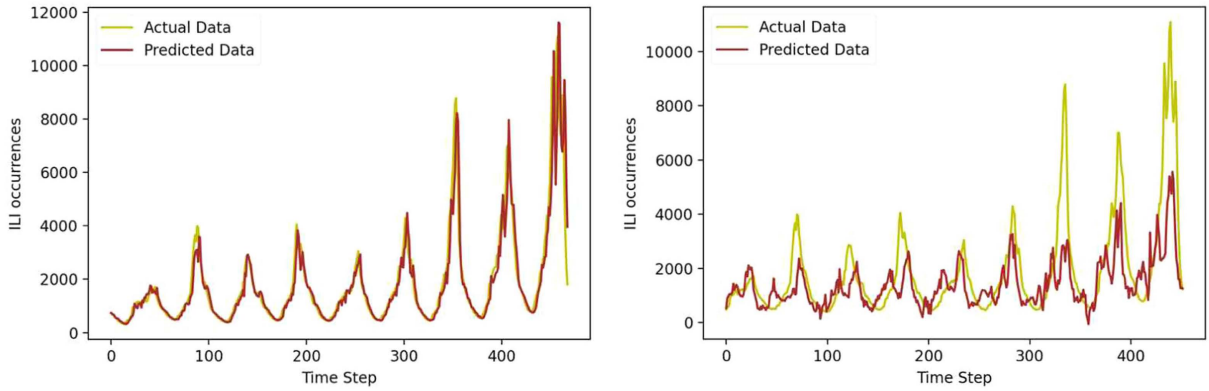


图20：使用大语言模型挖掘文本与时间序列数据来预测流感样疾病。横轴表示一个流感样疾病周期内的不同时间步，纵轴表示流感样疾病的发生次数。子图(a)比较了不同时间步提前两周的流感样疾病实际发生次数与预测发生次数，子图(b)则展示了提前二十周的相应结果。来源：转载自参考文献 [82]

文本和时间序列数据显著提高了传染病预测的准确性。这种预测主要有两种形式：一种是基于特定时间范围内的历史数据，直接预测未来的感染规模和传播轨迹；另一种则侧重于个体级预测，根据历史数据估计患者在时间t的感染状态。从人工智能的角度来看，后一种形式对应于分类任务，即模型根据过去的数据判断个体在时间t是否感染。如图21所示，Kang等人 [84] 展示了基于大语言模型的表示所带来的改进。可以看到，若不使用大语言模型表示而直接分析原始数据，表示向量很难分离。相比之下，经大语言模型表示后，表示向量可以得到有效分类。

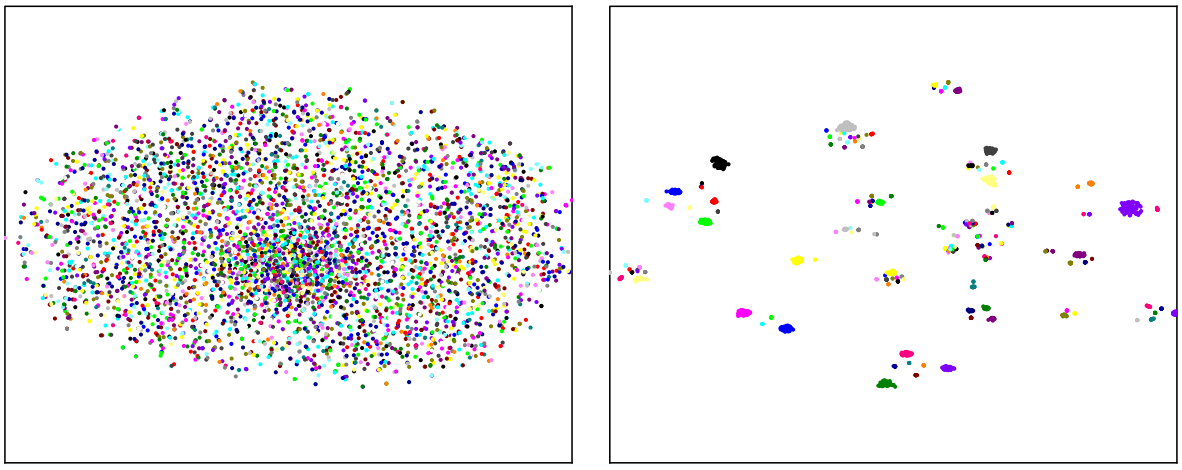


图21：疾病表示向量的可视化对比：(a) 不使用大语言模型表示；(b) 使用基于大语言模型的表示。未经过大语言模型增强的表示分布零散且相互重叠，各类别之间的边界不清晰，难以区分类别。相比之下，基于大语言模型的表示有了显著改善，在不同疾病类别中形成了彼此分离且定义清晰的簇。来源：转载自参考文献 [84]

除上述研究外，还有其他研究也有助于生物疫情的预测与管理。例如，Priya等人[133] 通过将传统流行病学模型与前沿人工智能技术相结合，开发了一种混合智能医疗系统，以提高疾病预测准确率并解决医疗可及性问题。研究人员利用大语言模型和自然语言处理技术，实时分析社交网络、新闻、政府公告及其他文本数据，从而增强分析能力。这使系统能够捕捉公众情绪和风险感知（例如恐慌信号），识别早期疫情关键词（例如有关症状的讨论），并追踪谣言传播，以协助制定干预策略。Jo等人[135] 探讨了为大语言模型驱动的聊天机器人添加长期记忆功能会如何影响用户关于健康的自我披露行为以及对聊天机器人的感知。这有助于通过鼓励个体在自然对话中披露健康信息来开展公共卫生监测。在与公共卫生机构合作时，市政官员会监控通话记录。如果聊天机器人检测到任何负面的健康信号（例如不进食、睡眠质量差、特定健康问题），或者有人多次未接听电话，官员便可联系该个体。McClymont等人[341]认为，人工智能和机器学习算法的进步增强了处理复杂大规模数据集的能力，加快了数据挖掘，并提升了预测模型的性能。

移动健康应用（包括人工智能聊天机器人）有望捕捉互联网用户的医疗服务寻求行为，以检测早期预警信号。

6. 结论与展望

传播动力学为理解复杂系统中的集体行为提供了基础视角。大语言模型的出现及其日益融入人类活动，推动传播动力学研究超越传统的基于结构和基于状态的表述，转向更丰富的表示方式，将文本语义、多模态信息和世界知识纳入其中。凭借大语言模型在语境理解和知识推理方面的强大能力，这些以往难以建模的因素如今得以有效捕捉，从而能够更准确地阐释传播背后的基本规律。此外，在网络空间中，如今海量谣言或错误信息由智能体传播，而这些智能体受少数大语言模型的引导或控制，因此形成的传播模式与大语言模型时代之前观察到的模式存在显著差异 [342]。这些变化推动了传播动力学领域新一轮研究。通过系统分析近期研究，本综述阐明了大语言模型如何被用于增强传播过程的建模、检测、预测和管理，评估了不同方法的有效性，并指出了关键挑战、尚未解释的现象和新兴趋势。总之，本综述提供了概念和方法层面的洞察力，不仅支持未来研究，也为致力于改善流行病、谣言及其他大规模传播事件的预测与管理的研究人员、机构和政策制定者提供实践指导。更广泛地说，这些研究有望推动各方就旨在保护公众和社会的政策干预措施的潜在成本与收益展开知情讨论。需要指出的是，将从传播模型中获得的洞察转化为有效且负责任的政策仍是一项艰巨任务，因为这一过程不可避免地涉及不确定性、取决于具体情境的风险，以及产生意外或不利后果的可能性。

展望未来，有几个方向尤其值得关注。

• 首先，大语言模型驱动的智能体建模与既有理论传播框架的整合仍有待深入探索。一项关键挑战在于，如何在基于语言的交互模式与可进行解析处理的动力学变量之间建立有原则的映射。

• 其次，大语言模型作为传播过程中的内生参与者而非外部工具，其作用会引入新的反馈路径，并引发关于传播系统稳定性、放大效应、共振和可控性的问题。随着大语言模型日益成为信息生成与传播的主体，它们所塑造的环境反过来也可能影响人类认知、决策和集体行为，并可能催生新兴的“机器文化” [35]。如何进一步将大语言模型所中介的信息生态系统与人类行为之间的这一反馈循环纳入更宏观的文化与社会动力学框架，以开展建模与分析，是理解未来人机共生通信系统不可或缺的研究方向。

• 第三，对大语言模型参与的传播过程进行评估，需要在测量 [343], 基准测试 [344；345； 88], 和鲁棒性 [346] 方面取得方法论进展。与具有明确转移规则的经典传播模型不同，基于大语言模型的动力学需要能够区分认知、内容驱动和结构性因素对观测结果所作贡献的框架。

• 第四，如何在大语言模型快速发展的背景下确保科学发现的鲁棒性和可复现性，是一个有待进一步探索的领域。随着模型架构、训练语料库和推理策略不断变化，基于大语言模型的传播分析结果可能会对特定模型版本、参数设置和实现细节变得高度敏感。这种不稳定性使跨研究比较更加复杂，也削弱了科学发现的可复现性。此外，关于传播的大语言模型驱动研究通常涉及复杂的感知、认知和生成过程，因此，基于固定规则的传统实验复现范式难以直接应用。为提高已报告结果的可靠性和长期有效性，未来研究可针对大语言模型驱动的传播研究建立标准化实验方案、版本控制实践和可复现性基准。

• 第五，超越成对传播，对高阶交互和集体效应进行建模，是另一个重要方向。现有的大多数传播模型都建立在节点之间的二元交互之上，而现实世界中的信息扩散往往受到群体讨论、协同行为和集体共振现象的影响。将此类高阶交互纳入涉及大语言模型的传播框架，或许能够更忠实地表示群体层面的协同、非线性放大效应及其对宏观传播模式的影响。

• 第六，大语言模型驱动的传播动力学中的可解释性问题尤为关键，需要持续关注。例如，在生物疫情建模与预测中，可解释性会直接影响世界卫生组织等机构制定干预策略。因此，模型预测不仅要让技术用户理解，也要让政策制定者和公众理解。然而，基于神经

网络架构构建的大语言模型常被视为“黑箱”，因为其内部决策过程在很大程度上难以为人类所理解。如果此类模型给出不准确的预测，误导干预决策并导致不良后果，就会引发对问责和可信性的严重担忧 [160]。

• 第七，大语言模型中的偏见问题也值得认真考虑。大语言模型的训练和微调依赖海量数据集，而这些数据集本身可能包含固有偏见。例如，许多生物医学数据集历来对某些人口群体代表不足，其中包括女性和少数族裔。这类偏见可能导致流行病预测与管理中的伦理错误或误读，进而损害公众对健康信息和决策的信任 [160]。

• 最后，大语言模型极易受到对抗攻击 [347；348； 349] 并且容易产生幻觉 [350；224]。一方面，在训练或部署阶段进行有针对性或无针对性的干扰，可能导致意外或误导性输出，严重损害传播动力学建模、检测和预测的准确性。因此，需要持续努力制定有效的防御策略，以应对此类攻击。另一方面，在流行病建模和分析过程中，如何避免模型幻觉造成错误判断，仍是一项需要持续关注的重要挑战。

总体而言，大语言模型与传播动力学的交汇正在塑造一种新的研究格局，其中内容、语义和行为被视为传播过程不可或缺的组成部分，而非外部修饰因素。推进这一方向需要统计物理学、网络科学、认知建模、计算社会科学和机器学习等领域的密切合作。这些跨学科努力不仅有助于构建表现力更强的传播模型，也有望加深我们对集体行为如何在日益互联的人机系统中形成、适应和演变的理解。

致谢

我们谨此感谢以下项目提供的部分资助：国家自然科学基金委员会（项目编号：62202320和62402331）、重庆医科大学未来医学青年创新计划（项目编号：W0150）、重庆市教育委员会科学技术研究计划项目（项目编号：KJQN202502828）、中央高校基本科研业务费（项目编号：YJ202429和SCU2024D012）、四川大学科学与工程内涵发展项目（项目编号：2020SCUNG129）、2024年重庆市教育委员会人文社会科学研究项目（项目编号：24SKGH048）、重庆市自然科学基金（项目编号：CSTB2025YITP-QCRCX0061）以及重庆市社会科学基金（项目编号：2025NDYB065）。

CRediT 作者贡献声明

Shuyu Jiang：撰写—审阅与编辑、撰写—初稿、方法学、指导、概念化。Hao Ren：撰写—初稿、方法学、调查。Yichang Gao：撰写—初稿、调查。Yi-Cheng Zhang：撰写—审阅与编辑、指导。Li Qi：撰写—审阅与编辑、概念化。Dayong Xiao：撰写—初稿、指导。Jie Fan：撰写—审阅与编辑、撰写—初稿、指导、概念化。Rui Tang：撰写—审阅与编辑、撰写—初稿、指导、概念化。Wei Wang：撰写—审阅与编辑、撰写—初稿、方法学、指导、概念化。Jie Fan(241031@cqmpc.edu.cn)、Rui Tang(tangrscu@scu.edu.cn) 和 Wei Wang(wwzqbx@hotmail.com) 为共同通讯作者。

参考文献

[1] S. H. Strogatz, 《探索复杂网络》, 《自然》410 (6825) (2001) 268–276。[2] J. M. Ottino, 《复杂系统工程学》, 《自然》427 (6973) (2004) 399–399。[3] Y. Bar-Yam, 《复杂系统动力学》, CRC出版社, 2019年。[4] E. O. Wilson, 《知识的统一：融会贯通》, 第31卷, Vintage出版社, 1999年。[5] A. Vespignani, 《复杂社会技术系统中的动力学过程建模》, 《自然·物理》8 (1) (2012) 32–39。[6] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, 《复杂网络：结构与动力学》, 《物理学报告》424 (4-5) (2006) 175–308。[7] M. De Domenico, C. Granell, M. A. Porter, A. Arenas, 《多层网络中传播过程的物理学》, 《自然·物理》12 (10) (2016) 901–906。[8] D. Notarmuzi, C. Castellano, A. Flammini, D. Mazzilli, F. Radicchi, 《社交媒体中信息传播的普适性、临界性与复杂性》, 《自然·通讯》13 (1) (2022) 1308。[9] D. Centola, M. Macy, 《复杂传染与弱关系的局限》, 《美国社会学杂志》113 (3) (2007) 702–734。

[10] D. Guilbeault, J. Becker, D. Centola, 《复杂传染：十年回顾》, 《社会系统中的复杂传播现象：现实世界社交网络中的影响与传染》（2018）3–25。[11] M. S. Mariani, F. Battiston, E.-Á. Horvát, G. Livan, F. Musciotto, D. Wang, 《成功背后的集体动力学》, 《自然·通讯》15 (1) (2024) 10701。[12] M. J. Keeling, P. Rohani, 《人类和动物传染病建模》, 普林斯顿大学出版社，2008年。[13] D. Centola, 《行为如何传播：复杂传染科学》, 第3卷，普林斯顿大学出版社（新泽西州普林斯顿），2018年。[14] H. W. Hethcote, 《传染病数学》, 《SIAM评论》42 (4) (2000) 599–653。[15] G. Ferraz de Arruda, A. Aleta, Y. Moreno, 《高阶网络上的传染动力学》, 《自然综述：物理》6 (8) (2024) 468–482。[16] W. Wang, Q.-H. Liu, J. Liang, Y. Hu, T. Zhou, 《复杂网络中的协同进化传播》, 《物理学报告》820 (2019) 1–51。[17] I. Z. Kiss, J. C. Miller, P. L. Simon 等, 《网络流行病数学》, 瑞士Cham：施普林格，598 (2017) (2017) 31。[18] W. Wang, Y. Nie, W. Li, T. Lin, M.-S. Shang, S. Su, Y. Tang, Y.-C. Zhang, G.-Q. Sun, 《高阶网络上的流行病传播》, 《物理学报告》1056 (2024) 1–70。[19] C. Gao, X. Lan, N. Li, Y. Yuan, J. Ding, Z. Zhou, F. Xu, Y. Li, 《大语言模型赋能的基于智能体的建模与仿真：综述与展望》, 《人文与社会科学通讯》11 (1) (2024) 1–24。[20] C. M. Macal, M. J. North, 《基于智能体的建模与仿真教程》, 载于：《冬季仿真会议论文集》, 2005年，IEEE, 2005年，第14–页。[21] J. T. Davis, N. Perra, Q. Zhang, Y. Moreno, A. Vespignani, 《结构化人群中信息传播的相变》, 《自然·物理》16 (5) (2020) 590–596。[22] J. Achiam, S. Adler, S. Agarwal, L. Ahmad, I. Akkaya, F. L. Aleman, D. Almeida, J. Altenschmidt, S. Altman, S. Anadkat 等, 《GPT-4技术报告》, arXiv预印本 arXiv:2303.08774 (2023)。[23] D. Guo, D. Yang, H. Zhang, J. Song, R. Zhang, R. Xu, Q. Zhu, S. Ma, P. Wang, X. Bi 等, 《DeepSeek-R1通过强化学习激励大语言模型进行推理》, 《自然》645 (2025年) 633–638。doi:10.1038/s41586-025-09422-z。[24] Y. Oh, S. Park, H. K. Byun, Y. Cho, I. J. Lee, J. S. Kim, J. C. Ye, 《放射肿瘤学中的大语言模型驱动多模态靶区勾画》, 《自然·通讯》15 (1) (2024) 9186。[25] B. Romera-Paredes, M. Barekatain, A. Novikov, M. Balog, M. P. Kumar, E. Dupont, F. J. Ruiz, J. S. Ellenberg, P. Wang, O. Fawzi 等, 《利用大语言模型通过程序搜索实现数学发现》, 《自然》625 (7995) (2024) 468–475。[26] P. Kumar, 《大语言模型（LLM）：综述、技术框架与未来挑战》, 《人工智能评论》57 (10) (2024) 260。[27] J. Gui, T. Chen, J. Zhang, Q. Cao, Z. Sun, H. Luo, D. Tao, 《自监督学习综述：算法、应用与未来趋势》, 《IEEE模式分析与机器智能汇刊》46 (12) (2024) 9052–9071。[28] A. Chowdhery, S. Narang, J. Devlin, M. Bosma, G. Mishra, A. Roberts, P. Barham, H. W. Chung, C. Sutton, S. Gehrmann 等, 《PaLM：通过路径实现语言建模的标度扩展》, 《机器学习研究杂志》24 (240) (2023) 1–113。[29] T. Eloundou, S. Manning, P. Mishkin, D. Rock, 《GPT就是GPT：大语言模型的劳动力市场影响潜力》, 《科学》384 (6702) (2024) 1306–1308。[30] E. M. Fenoaltea, D. Mazzilli, A. Patelli, A. Sbardella, A. Tacchella, A. Zaccaria, M. Trombetti, L. Pietronero, 《追踪资金流向：一种基于初创企业的衡量指标，用于评估各职业、行业和地区的人工智能暴露度》, arXiv预印本 arXiv:2412.04924 (2024)。[31] K. Swanson, W. Wu, N. L. Bulaong, J. E. Pak, J. Zou, 《人工智能智能体虚拟实验室设计出新型SARS-CoV-2纳米抗体》, 《自然》（2025年）1–3。[32] G. De Vito, 《评估基于物联网和大语言模型技术构建的医疗保健软件》, 载于：第28届软件工程评估与测量国际会议论文集，2024年，第476–481页。[33] Y. Kong, Y. Nie, X. Dong, J. M. Mulvey, H. V. Poor, Q. Wen, S. Zohren, 《面向金融与投资管理的大语言模型：应用与基准》, 《投资组合管理杂志》51 (2) (2024)。[34] B. Yang, Y. Guo, L. Xu, Z. Yan, H. Chen, G. Xing, X. Jiang, 《Socialmind：基于大语言模型的主动式增强现实社交辅助系统，具备类人感知能力，可实现现场实时交互》, 《ACM交互、移动、可穿戴与泛在技术汇刊》9 (1) (2025) 1–30。

[35] L. Brinkmann、F. Baumann、J.-F. Bonnefon、M. Derex、T. F. Müller、A.-M. Nussberger、A. Czaplicka、A. Acerbi、T. L. Griffiths、J. Henrich等，《机器文化》，《自然·人类行为》7（11）（2023）1855–1868。[36] J. R. Anthis、R. Liu、S. M. Richardson、A. C. Kozlowski、B. Koch、E. Brynjolfsson、J. Evans、M. S. Bernstein，《立场：大语言模型社会模拟是一种有前景的研究方法》，载于：第四十二届国际机器学习会议立场论文专题，2025年。[37] A. Kozlowski、J. Evans，《模拟受试者：用于社会主体与社会互动的人工智能替身的前景与风险》（2024）。[38] D. T. Schroeder、M. Cha、A. Baronchelli、N. Bostrom、N. A. Christakis、D. Garcia、A. Goldenberg、Y. Kyrychenko、K. Leyton-Brown、N. Lutz等，《恶意人工智能群体如何威胁民主》，arXiv预印本 arXiv:2506.06299（2025）。[39] L. Li、J. Li、C. Chen、F. Gui、H. Yang、C. Yu、Z. Wang、J. Cai、J. A. Zhou、B. Shen等，《政治大语言模型：政治学中的大语言模型》，arXiv预印本 arXiv:2412.06864（2024）。[40] Y. Lu、A. Aleta、C. Du、L. Shi、Y. Moreno，《用于复杂系统研究的大语言模型与生成式基于智能体模型》，《生命物理学评论》51（2024）283–293。[41] M. Tsvetkova、T. Yasseri、N. Pescetelli、T. Werner，《人类与机器的新社会学》，《自然·人类行为》8（10）（2024）1864–1876。[42] D. Pedreschi、L. Pappalardo、E. Ferragina、R. Baeza-Yates、A.-L. Barabási、F. Dignum、V. Dignum、T. Eliassi-Rad、F. Giannotti、J. Kertész等，《人机协同进化》，《人工智能》339（2025）104244。[43] H. Bai、J. G. Voelkel、S. Muldowney、J. C. Eichstaedt、R. Willer、M. S. Bernstein，《大语言模型生成的信息能够说服人类接受政策议题上的观点》，《自然·通讯》16（1）（2025）6037。[44] K. Radivojevic、N. Clark、P. Brenner，《我们身边的大语言模型：生成式人工智能参与数字话语》，载于：AAAI研讨会系列论文集，第3卷，2024年，第209–218页。[45] Z. Xi、W. Chen、X. Guo、W. He、Y. Ding、B. Hong、M. Zhang、J. Wang、S. Jin、E. Zhou等，《基于大语言模型的智能体的兴起与潜力：综述》，《中国科学：信息科学》68（2）（2025）121101。[46] J. S. Park、J. O'Brien、C. J. Cai、M. R. Morris、P. Liang、M. S. Bernstein，《生成式智能体：人类行为的交互式拟像》，载于：第36届ACM用户界面软件与技术研讨会论文集，2023年，第1–22页。[47] X. Shen、Y. Liu、Y. Dai、Y. Wang、R. Miao、Y. Tan、S. Pan、X. Wang，《理解通信拓扑对基于大语言模型的多智能体系统中信息传播的影响》，载于：2025年自然语言处理实证方法会议论文集，2025年，第12357–12372页。[48] J. Piao、Z. Lu、C. Gao、Y. Li，《社交机器人遇上大语言模型：政治偏见与受社会学习启发的缓解策略》，载于：2025年ACM网络会议论文集，2025年，第5202–5211页。[49] D. Smilkov、C. A. Hidalgo、L. Kocarev，《超越网络结构：异质易感性如何调节流行病传播》，《科学报告》4（1）（2014）4795。[50] R. G. Seymour、T. Kypraios、P. D. O'Neill，《异质混合传染病模型的贝叶斯非参数推断》，《美国国家科学院院刊》119（10）（2022）e2118425119。[51] S. Aral、P. S. Dhillon，《基于实证影响力模型的社会影响力最大化》，《自然·人类行为》2（6）（2018）375–382。[52] M. Gomez-Rodriguez、J. Leskovec、A. Krause，《推断扩散与影响网络》，《ACM知识发现与数据》5（4）（2012）1–37。[53] T. W. Valente，《创新扩散中的社会网络阈值》，《社会网络》18（1）（1996）69–89。[54] R. Tanase、R. Algesheimer、M. S. Mariani，《将行为实验发现整合到动力学模型中以指导社会变革干预》，arXiv预印本 arXiv:2405.13224（2024）。[55] T. P. Peacock、L. Moncla、G. Dudas、D. VanInsberghe、K. Sukhova、J. O. Lloyd-Smith、M. Worobey、A. C. Lowen、M. I. Nelson，《哺乳动物中的全球H5N1流感泛兽疫》，《自然》637（8045）（2025）304–313。[56] M. Budia-Silva、T. Kostyanev、S. Ayala-Montaño、J. Bravo-Ferrer Acosta、M. Garcia-Castillo、R. Cantón、H. Goossens、J. Rodriguez-Baño、H. Grundmann、S. Reuter，《耐碳青霉烯类肺炎克雷伯菌在欧洲的国际与区域传播》，《自然·通讯》15（1）（2024）5092。[57] H. Ni、X. Cai、J. Ren、T. Dai、J. Zhou、J. Lin、L. Wang、L. Wang、S. Pei、Y. Yao等，《中国登革热的流行病学特征与传播动力学》，《自然·通讯》15（1）（2024）8060。[58] M. Aguiar、V. Anam、K. B. Blyuss、C. D. S. Estadilla、B. V. Guerrero、D. Knopoff、B. W. Kooi、A. K. Srivastav、V. Steindorf、N. Stollenwerk，《登革热流行病学的数学模型：十年系统综述》，《生命物理学评论》40（2022）65–92。[59] S. F. Nasrin、C. Rajivganthi，《具有非线性发病率函数的随机埃博拉模型的动力学分析》，《非线性科学期刊》35（1）（2025）33。

[60] A. K. Keita、F. R. Koundouno、M. Faye、A. Düx、J. Hinzmann、H. Diallo、A. Ayouba、F. Le Marcis、B. Soropogui、K. Ifono等，《2021年几内亚埃博拉病毒复苏提示疫情暴发的新范式》，《自然》597（7877）（2021）539–543。[61] R. Williams、N. Hosseinichimeh、A. Majumdar、N. Ghaffarzadegan，《使用生成式智能体进行流行病建模》，arXiv预印本 arXiv:2307.04986（2023）。[62] X. Li、Y. Xu、Y. Zhang、E. C. Malthouse，《不同网络结构下用于新闻传播的大语言模型驱动多智能体仿真》，arXiv预印本 arXiv:2410.13909（2024）。[63] Y. Liu、X. Chen、X. Zhang、X. Gao、J. Zhang、R. Yan，《从怀疑到接受：模拟对假新闻的态度动态》，载于：IJCAI，2024年。[64] J. Ji、R. Lei、J. Bi、Z. Wei、X. Chen、Y. Lin、X. Pan、Y. Li、B. Ding，《基于大语言模型的多智能体系统是可扩展的图生成模型》，载于：计算语言学协会论文集：ACL 2025，2025年，第1492–1523页。[65] C. Wang、Z. Liu、D. Yang、X. Chen，《解码回音室：揭示社交网络极化的大语言模型驱动仿真》，载于：第31届计算语言学国际会议论文集，2025年，第3913–3923页。[66] Y.-S. Chuang、A. Goyal、N. Harlalka、S. Suresh、R. Hawkins、S. Yang、D. Shah、J. Hu、T. Rogers，《使用基于大语言模型的智能体网络模拟观点动力学》，载于：计算语言学协会北美分会论文集：NAACL 2024，2024年，第3326–3346页。[67] Y. Liu、Z. Song、J. Zhang、X. Zhang、X. Chen、R. Yan，《渐进式欺骗：使用大语言模型智能体模拟从真实新闻到假新闻的演变》，载于：2025年自然语言处理实证方法会议论文集，2025年，第26187–26203页。[68] W. Ma、C. Deng、A. Moossavi、L. Wang、S. Vosoughi、D. Yang，《模拟错误信息易感性（SMISTS）：利用大语言模型仿真增强错误信息研究》，载于：计算语言学协会论文集：ACL 2024，2024年，第2774–2788页。[69] X. Mou、Z. Wei、X.-J. Huang，《揭示真相并促进变革：迈向基于智能体的大规模社会运动仿真》，载于：计算语言学协会论文集：ACL 2024，2024年，第4789–4809页。[70] J. Yao、H. Zhang、J. Ou、D. Zuo、Z. Yang、Z. Dong，《社会观点预测利用动力学方程与基于大语言模型的智能体融合》，《科学报告》15（1）（2025）15472。[71] Y. Liu、W. Liu、X. Gu、W. Wang、J. Luo、Y. Zhang，《RumorSphere：百万规模基于智能体的谣言传播动态仿真框架》，arXiv预印本 arXiv:2509.02172（2025）。[72] Z. Cheng、Z. Lin、Y. Yang、Z. Wei、S. Chen，《基于大语言模型的多智能体建模对社交媒体事件动态的交互式仿真与可视化分析》，《可视化信息学》（2025）100260。[73] W. Jiang、T. Chen、X. Gao、W. Zhang、L. Cui、H. Yin，《流行病学信息网络用于稳健的谣言检测》，载于：2025年ACM网络会议论文集，2025年，第3618–3627页。[74] B. Qiao、K. Li、W. Zhou、S. Li、Q. Lu、S. Hu，《BotSim：大语言模型驱动的恶意社交僵尸网络仿真》，载于：AAAI人工智能会议论文集，第39卷，2025年，第14377–14385页。[75] G. De Marzo、L. Pietronero、D. Garcia，《大语言模型社会互动中无标度网络的涌现》，arXiv预印本 arXiv:2312.06619（2023）。[76] S. Chang、A. Chaszczewicz、E. Wang、M. Josifovska、E. Pierson、J. Leskovec，《大语言模型生成结构逼真的社交网络，但高估了政治同质性》，载于：国际AAAI网络与社交媒体会议论文集，第19卷，2025年，第341–371页。[77] M. Papachristou、Y. Yuan，《多大语言模型之间的网络形成与动态》，《PNAS Nexus》4（12）（2025）pgaf317。[78] Z. Lu、G. Lim、M. Yin，《理解大语言模型驱动的对抗性社会影响对在线信息传播的影响》，载于：计算机系统中的人为因素会议扩展摘要论文集，2025年，第1–7页。[79] C. Bandara，《作为虚假信息的幻觉：大语言模型在放大阴谋论和假新闻中的作用》，《应用网络安全分析、情报与决策系统期刊》14（12）（2024）65–76。[80] H. D. L. Villaplana、J. Kwak、M. H. Lees、H. Li、W. Cai，《生成式人工智能在流行病建模中的应用》，载于：2024年冬季仿真会议（WSC），IEEE，2024年，第2727–2738页。[81] C. Gong、R. Sun、Y. Zheng、J. Zhang、T. Gu、L. Pan、L. Lv，《EpiLLM：释放大语言模型在疫情预测中的潜力》，arXiv预印本 arXiv:2505.12738（2025）。[82] J. Moon、J. Shim、E. Kim、E. Hwang，《MIFlu：基于大语言模型的多模态流感预测方案》，《IEEE生物医学与健康信息学汇刊》（2025）。[83] H. Du、Y. Zhao、J. Zhao、S. Xu、X. Lin、Y. Chen、L. M. Gardner、H. F. Yang，《利用大语言模型推进实时传染病预测》，《自然·计算科学》（2025）1–14。[84] Y. Kang、M. Yang、Y. Peng、J. Cai、L. Zhao、Z. Gao、N. Li、B. Pu，《LLM-DG：通过患者间与患者内建模利用大语言模型增强疾病预测》，《信息融合》121（2025）103145。

[85] K. O. Kwok, T. Huynh, W. I. Wei, S. Y. Wong, S. Riley, A. Tang, 将大语言模型用于传染病传播建模以做好公共卫生应急准备，《计算与结构生物技术期刊》23 (2024年) 3254–3257。[86] I. Zaslavsky, J. Lei, R. Graham, M. S. Handcock, E. Aronoff-Spencer, 利用大语言模型增强空间细分模拟，载于：《2024年第七届数学与统计学国际会议论文集》，2024年，第14–18页。[87] J. Wang, R. Jiang, C. Yang, Z. Wu, M. Onizuka, R. Shibasaki, N. Koshizuka, C. Xiao, 作为城市居民的大语言模型：用于个人出行生成的大语言模型智能体框架，《神经信息处理系统进展》37 (2024年) 124547–124574。[88] Z. Wu, R. Peng, T. Ito, C. Xiao, 基于大语言模型的社会模拟需要边界，arXiv预印本 arXiv:2506.19806 (2025年)。[89] J. Nudo, M. E. Pandolfo, E. Loru, M. Samory, M. Cinelli, W. Quattrociocchi, 大语言模型社会智能体中的生成式夸大：一致性、偏见与毒性，《在线社交网络与媒体》51 (2026) 100344。[90] J. Piao, Y. Yan, J. Zhang, N. Li, J. Yan, X. Lan, Z. Lu, Z. Zheng, J. Y. Wang, D. Zhou, 等，AgentSociety：大语言模型驱动的生成式智能体大规模模拟，助力深入理解人类行为与社会，arXiv预印本 arXiv:2502.08691 (2025年)。[91] H. Guan, J. He, L. Fan, Z. Ren, S. He, X. Yu, Y. Chen, S. Zheng, T.-Y. Liu, Z. Liu, 利用十亿个智能体建模地球尺度的类人社会，arXiv预印本 arXiv:2506.12078 (2025年)。[92] P. Salem, R. Sim, C. Olsen, P. Saxena, R. Barcelos, Y. Ding, TinyTroupe：大语言模型驱动的多智能体角色设定模拟工具包，arXiv预印本 arXiv:2507.09788 (2025年)。[93] B. Hu, Q. Sheng, J. Cao, Y. Shi, Y. Li, D. Wang, P. Qi，坏演员，好顾问：探索大语言模型在假新闻检测中的作用，载于：AAAI人工智能会议论文集，第38卷，2024年，第22105–22113页。[94] Q. Nan, Q. Sheng, J. Cao, B. Hu, D. Wang, J. Li，让沉默发声：利用大语言模型生成的评论增强假新闻检测，载于：第33届ACM信息与知识管理国际会议论文集，2024年，第1732–1742页。[95] J. Ma, L. Hu, R. Li, W. Fu，LOCAL：基于大语言模型的多智能体进行逻辑与因果事实核查，载于：2025年ACM网络会议论文集，2025年，第1614–1625页。[96] R. Zong, Y. Zhang, D. Wang，赋能大语言模型综合人工智能与人类智能，以实现社交媒体上可解释的公共卫生错误信息检测，载于：国际AAAI网络与社交媒体会议论文集，第19卷，2025年，第2334–2348页。[97] B. Xie, X. Ma, X. Shan, A. Beheshti, J. Yang, H. Fan, J. Wu，面向假新闻检测的多知识与大语言模型启发式异构图神经网络，《IEEE计算社会系统汇刊》(2024)。[98] C. Zhang, Z. Feng, Z. Zhang, J. Qiang, G. Xu, Y. Li，大语言模型的幻觉是否可用？用于假新闻检测的大语言模型负向推理，载于：AAAI人工智能会议论文集，第39卷，2025年，第1031–1039页。[99] J. Lucas, A. Uchendu, M. Yamashita, J. Lee, S. Rohatgi, D. Lee，以火攻火：大语言模型在编造和检测隐蔽虚假信息中的双重作用，载于：2023年自然语言处理实证方法会议论文集，计算语言学协会，2023年。[100] X. Ma, Y. Zhang, K. Ding, J. Yang, J. Wu, H. Fan，基于大语言模型增强语义挖掘的假新闻检测，载于：2024年自然语言处理实证方法会议论文集，2024年，第508–521页。[101] R. Yang, J. Ma, W. Gao, H. Lin，利用大语言模型增强的多示例学习，结合社会语境信息进行谣言与立场联合检测，《ACM智能系统与技术汇刊》16 (3) (2025) 1–27。[102] K. Krykoniuk, C. Hopkin-King, S. G. Roberts，利用大语言模型通过话语偏离度量检测虚假信息 (2024)。[103] L. Hu, Z. Wang, J. Zhu, Y. Hu, X. Wang，MAGE-fend：借助大语言模型专业知识指导的多模态自适应融合，用于短视频平台上的假新闻检测，Knowledge-Based Systems (2025) 114298。[104] J. Chen, L. Liu, T. Lan, F. Zhou, X. Zhang，你只需查询两次：从双重视角通过证据评估进行多模态谣言检测，载于：第31届计算语言学国际会议论文集，2025年，第3415–3427页。[105] Y. Bai, K. Fu，基于大语言模型并结合RAG事实核查的假新闻检测框架，载于：2024年IEEE国际大数据会议（BigData），IEEE，2024年，第8617–8619页。[106] X. Zeng, D. La Barbera, K. Roitero, A. Zubiaga, S. Mizzaro，结合大语言模型与众包进行人机混合错误信息检测，载于：第47届ACM SIGIR信息检索研究与发展国际会议论文集，2024年，第2332–2336页。[107] Y. Liu, Y. Liu, X. Zhang, X. Chen, R. Yan，辩论让真相更加清晰！利用大语言模型的多智能体系统揭露假新闻，载于：第48届ACM SIGIR信息检索研究与发展国际会议论文集，2025年，第504–514页。

[108] X. Li, Y. Zhang, E. C. Malthouse，基于大语言模型智能体的事实核查与假新闻检测方法，载于：ECAI 2024，IOS出版社，2024年，第2572–2579页。[109] A. Xie, F. Zhu, J. Han, S. Hu，通过大语言模型整合开放域知识以进行多模态假新闻检测，载于：2024年第27届计算机支持的协同工作与设计国际会议（CSCWD），IEEE，2024年，第1917–1922页。[110] 评估生成式智能体在众包事实核查中的潜力，《在线社交网络与媒体》48 (2025) 100326。[111] H. Wan, S. Feng, Z. Tan, H. Wang, Y. Tsvetkov, M. Luo，DELL：为基于大语言模型的错误信息检测生成反应与解释，载于：计算语言学协会年会 ACL 2024 论文集，2024年，第2637–2667页。[112] G. Li, D. Hu, Z. Liu, X. Zhang, H. Lyu，利用大语言模型语义重排与异构图自编码器增强谣言检测，载于：第31届计算语言学国际会议论文集，2025年，第8557–8572页。[113] Y. Zeng, X. Ding, B. Cai, T. Liu, B. Qin，探索大语言模型在社交媒体有效谣言检测中的应用，载于：2025年美洲计算语言学协会分会会议：人类语言技术论文集（第一卷：长篇论文），2025年，第2537–2552页。[114] S. Hu, J. Hu, H. Zhang，结合大语言模型与全局标签传播进行多模态假新闻检测，载于：W. Che, J. Nabende, E. Shutova, M. T. Pilehvar（编），第63届计算语言学协会年会论文集（第一卷：长篇论文），计算语言学协会，维也纳，奥地利，2025年，第1426–1440页。[115] M. S. Deiner, R. Y. Deiner, C. Fathy, N. A. Deiner, V. Hristidis, S. D. McLeod, T. J. Bukowski, T. Doan, G. D. Seitzman, T. M. Lietman, 等，利用大语言模型对社交媒体上关于结膜炎暴发的合成及现实世界帖子中的流行病学特征进行分类：信息流行病学研究，《医学互联网研究杂志》27 (2025) e65226。[116] M. S. Deiner, N. A. Deiner, V. Hristidis, S. D. McLeod, T. Doan, T. M. Lietman, T. C. Porco，利用大语言模型根据推文内容评估流行病暴发的可能性：信息流行病学研究，《医学互联网研究杂志》26 (2024) e49139。[117] J. Kaur, Z. A. Butt，人工智能驱动的疫情情报：疫情暴发检测与响应的未来，Frontiers in Artificial Intelligence 8 (2025) 1645467。[118] J. Xie, Z. Zhang, S. Zeng, J. Hilliard, G. An, X. Tang, L. Jiang, Y. Yu, X. Wan, D. Xu, 等，利用大语言模型进行传染病监测——使用网络服务监测基于自我报告推文的新冠-19模式：内容分析，《医学互联网研究杂志》27 (1) (2025) e63190。[119] S. Consoli, P. Markov, N. I. Stilianakis, L. Bertolini, A. P. Gallardo, M. Ceresa，利用大语言模型为基于事件的监测抽取疫情信息，载于：信息与通信技术国际大会，Springer Nature Singapore，新加坡，2024年，第241–252页。[120] F. Sufi，使用大语言模型分析新冠肺炎主题的一种创新方法，《经济与技术杂志》(2024)。[121] X. Wang, M. Feng, J. Qiu, J. Gu, J. Zhao，从新闻到预测：在基于大语言模型的时间序列预测中整合事件分析与反思，《神经信息处理系统进展》37 (2024) 58118–58153。[122] T. Zhong, J. Zhang, Z. Cheng, F. Zhou, X. Chen，通过级联检索上下文学习进行信息扩散预测，载于：第47届ACM SIGIR信息检索研究与发展国际会议论文集，2024年，第2472–2476页。[123] W. Shang，利用大语言模型预测社交网络中的扩散：一种融合注意力感知模型，载于：2025年IEEE第41届数据工程国际会议（ICDE），IEEE计算机学会，2025年，第4685–4689页。[124] T. Chen, J. Huang, X. Wu, C. Shao，利用大语言模型预测社交媒体热度：将元数据转换为语义增强与情境化文本，《IEEE Access》(2024)。[125] Y. Zheng, C. Gong, R. Sun, J. Zhang, L. Pan, L. Lv，AutoCas：利用大语言模型进行社交网络自回归级联预测，arXiv预印本 arXiv:2502.18040 (2025年)。[126] J. Gao, X. Wang, T. Gan, J. Yin, C. Luo, L. Nie，社会情境感知的社区级传播预测，载于：第48届ACM SIGIR信息检索研究与发展国际会议论文集，2025年，第1995–2005页。[127] W.-J. Sun, X. F. Liu，利用大语言模型学习用户时序特征以预测转发，《计算机、材料与连续体》82 (3) (2025)。[128] C. Gao, X. Lan, Z. Lu, J. Mao, J. Piao, H. Wang, D. Jin, Y. Li，S3：由大语言模型赋能智能体的社交网络模拟系统，arXiv预印本 arXiv:2307.14984 (2023年)。[129] G. Rossetti, M. Stella, R. Cazabet, K. Abramski, E. Cau, S. Citraro, A. Failla, R. Improta, V. Morini, V. Pansanella，Y Social：大语言模型驱动的社交媒体数字孪生，arXiv预印本 arXiv:2408.00818 (2024年)。[130] L. Zhang, Y. Hu, W. Li, Q. Bai, P. Nand，LLM-AIDSim：用于社交网络中基于智能体的影响力扩散模拟的大语言模型增强系统，Systems 13 (1) (2025) 29。

[131] G. Liu, V. Le, S. Rahman, E. Kreiss, M. Ghassemi, S. Gabriel, 《MOSAIC：面向多智能体模拟中内容传播与监管的社交人工智能建模》，arXiv预印本 arXiv:2504.07830（2025年）。[132] F. Saeed, M. Aldosari, I. B. Arpinar, J. A. Miller, 《LLM4Cast：用于病毒性疾病预测的改造版大语言模型》，载于：2024 IEEE国际大数据会议（BigData），IEEE，2024，页码1428–1433。[133] E. Priya, D. Kumar, K. Jayachandiran, R. Shamprakash, 《采用流行病学建模的智能医疗助理》，载于：2024年电力、能源、控制与传输系统国际会议（ICPECTS），IEEE，2024，页码1–5。[134] M. Almaliki, A. M. Almars, K. O. Aljuhani, E.-S. Atlam, 《结合大语言模型与说服技术的优势来应对网络仇恨》，《计算机》14（5）（2025年）173。[135] E. Jo, Y. Jeong, S. Park, D. A. Epstein, Y.-H. Kim, 《了解长期记忆对用于公共卫生干预的大语言模型驱动聊天机器人自我披露的影响》，载于：2024年计算机系统中的人为因素会议论文集，2024，页码1–21。[136]A. Aleta, A. S. Teixeira, G. F. de Arruda, A. Baronchelli, A. Barrat, J. Kertész, A. Díaz-Guilera, O. Ar- time, M. Starnini, G. Petri, 等, 《多层网络科学：理论、方法与应用》，arXiv预印本 arXiv:2511.23371（2025年）。[137] R. Tang, Z. Yong, S. Jiang, X. Chen, Y. Liu, Y.-C. Zhang, G.-Q. Sun, W. Wang, 《网络对齐》，《物理学报告》1107（2025年）1–45。[138] M. Pósfai, A.-L. Barabási, 《网络科学》，第3卷，剑桥大学出版社，英国剑桥，2016。[139] A.-L. Barabási, R. Albert, 《随机网络中标度的涌现》，《科学》286（5439）（1999）509–512。[140] I. Varga, 《通过广告模拟比较网络拓扑》，载于：复杂性、未来信息系统与风险国际会议，第2卷，SCITEPRESS，2017，页码17–22。[141] S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, M. Zanin, 《多层网络的结构与动力学》，《物理学报告》544（1）（2014）1–122。[142] P. Holme, J. Saramäki, 《时序网络》，《物理学报告》519（3）（2012）97–125。[143] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, G. Petri, 《超越两两相互作用的网络：结构与动力学》，《物理学报告》874（2020年）1–92。[144] R. Lambiotte, M. Rosvall, I. Scholtes, 《从网络到复杂系统的最优高阶模型》，《自然·物理》15（4）（2019年）313–320。[145] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, 等, 《复杂系统中高阶交互的物理学》，《自然·物理》17（10）（2021年）1093–1098。[146] C. Qian, D. Zhao, B. Zhang, M. Zhong, J. Han, S. Li, H. Peng, W. Wang, 《相互依赖的高阶网络抵御级联失效的一对多鲁棒性》，《IEEE可靠性汇刊》（2025年）。[147] H. Peng, G. Pan, B. Zhang, C. Qian, M. Zhong, S. Li, J. Han, D. Zhao, W. Wang, 《高阶社交网络中由哨兵触发的错误信息遏制竞争动力学策略》，《非线性动力学》113（25）（2025年）35167–35181。[148] Y. Gao, J. Li, F. Gao, W. Wang, 《多种病原体在高阶网络上的协同演化》，《混沌、孤子与分形》202（2026年）117588。[149] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, 《复杂网络中的流行病过程》，《现代物理评论》87（3）（2015）925–979。[150] R. Ghosh, K. Lerman, 《网络上级联的定量分析框架》，载于：第四届ACM网络搜索与数据挖掘国际会议论文集，2011，页码665–674。[151] W. Wang, M. Tang, H. E. Stanley, L. A. Braunstein, 《复杂网络中流行病传播理论方法的统一》，《物理学进展报告》80（3）（2017）036603。[152] D. Centola, 《影响者、反噬效应与边缘的力量》，《个人网络：自我中心分析的经典读物与新方向》51（2021）73。[153] C. Castellano, S. Fortunato, V. Loreto, 《社会动力学的统计物理学》，《现代物理评论》81（2）（2009）591–646。[154] Z.-K. Zhang, C. Liu, X.-X. Zhan, X. Lu, C.-X. Zhang, Y.-C. Zhang, 《信息扩散动力学及其在复杂网络中的应用》，《物理学报告》651（2016）1–34。[155] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, I. Polosukhin, 《注意力就是一切》，《神经信息处理系统进展》第30卷（2017）。[156] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, 等, 《语言模型是少样本学习器》，《神经信息处理系统进展》第33卷（2020年）1877–1901。

[157] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, 等, 《LLaMA：开放且高效的基础语言模型》, arXiv预印本 arXiv:2302.13971（2023年）。[158] P. Tavana, M. ZareiNejad, 《生成式人工智能在大流行病预测中的应用》, 载于：《医疗系统中的生成式人工智能》, 施普林格, 2025, 页码35–66。[159] G. Bizel-Bizellot, S. Galmiche, B. Lelandais, T. Charmet, L. Coudeville, A. Fontanet, C. Zimmer, 《利用大语言模型从自由文本中提取新冠病毒传播情境》, 《自然·通讯》16（1）（2025年）5836。[160] M. U. Kraemer, J. L.-H. Tsui, S. Y. Chang, S. Lytras, M. P. Khurana, S. Vanderslott, S. Bajaj, N. Scheidwasser, J. L. Curran-Sebastian, E. Semenova, 等, 《用于传染病流行建模的人工智能》, 《自然》638（8051）（2025年）623–635。[161] H. Li, R. C. Gerkin, A. Bakke, R. Norel, G. Cecchi, C. Laudamiel, M. Y. Niv, K. Ohla, J. E. Hayes, V. Parma, 等, 《根据自我报告的化学感觉描述以文本为基础预测新冠病毒诊断结果》, 《通讯·医学》3（1）（2023年）104。[162] G. Hao, J. Wu, Q. Pan, R. Morello, 《量化大语言模型幻觉在复杂适应性社会网络中传播的不确定性》, 《科学报告》14（1）（2024年）16375。[163] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, 《BERT：用于语言理解的深度双向变换器预训练》, 载于：2019年计算语言学协会北美分会会议论文集：人类语言技术, 第1卷（长短论文）, 2019, 页码4171–4186。[164] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, 等, 《通过人类反馈训练语言模型遵循指令》, 《神经信息处理系统进展》第35卷（2022年）27730–27744。[165] Y. Wang, Y. Kordi, S. Mishra, A. Liu, N. A. Smith, D. Khashabi, H. Hajishirzi, 《Self-Instruct：利用自生成指令对齐语言模型》, 载于：第61届计算语言学协会年会论文集（第1卷：长篇论文）, 2023, 页码13484–13508。[166] H. W. Chung, L. Hou, S. Longpre, B. Zoph, Y. Tay, W. Fedus, Y. Li, X. Wang, M. Dehghani, S. Brahma, 等, 《指令微调语言模型的标度》, 《机器学习研究杂志》25（70）（2024年）1–53。[167] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, 等, 《通过人类反馈训练语言模型遵循指令》, 《神经信息处理系统进展》第35卷（2022年）27730–27744。[168] H. Lee, S. Phatale, H. Mansoor, T. Mesnard, J. Ferret, K. Lu, C. Bishop, E. Hall, V. Carbune, A. Rastogi, 等, 《RLAIF与RLHF：借助人工智能反馈扩展基于人类反馈的强化学习》, 载于：第41届国际机器学习会议论文集, 2024, 页码26874–26901。[169] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov, 《近端策略优化算法》, arXiv预印本 arXiv:1707.06347（2017）。[170] S. Bushaj, X. Yin, A. Beqiri, D. Andrews, ̇I. E. Büyüktahtakın, 《一种用于疫情控制优化的仿真深度强化学习（SIRL）方法》, 《运筹学年刊》328（1）（2023年）245–277。[171] R. Rakhshandehroo, D. Coombs, 《空间疫情模拟的奖励工程：用于个体行为学习的强化学习平台》, arXiv预印本 arXiv:2511.18000（2025年）。[172] F. Liu, T. Zhu, X. Wu, B. Yang, C. You, C. Wang, L. Lu, Z. Liu, Y. Zheng, X. Sun, 等, 《面向未来大流行病的医学多模态大语言模型》, NPJ Digital Medicine 6（1）（2023年）226。[173] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, 等, 《一幅图像胜过16×16个词：用于大规模图像识别的变换器》, 载于：国际学习表征会议论文集, 2021。[174] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner, G. Monfardini, 《图神经网络模型》, 《IEEE神经网络汇刊》20（1）（2008）61–80。[175] Z. Du, Y. Qian, X. Liu, M. Ding, J. Qiu, Z. Yang, J. Tang, 《GLM：通过自回归空白填充进行通用语言模型预训练》, 载于：第60届计算语言学协会年会论文集（第1卷：长篇论文）, 2022, 页码320–335。[176] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P. J. Liu, 《探索统一文本到文本变换器的迁移学习极限》, 《机器学习研究杂志》21（140）（2020年）1–67。[177] V. Sanh, A. Webson, C. Raffel, S. H. Bach, L. Sutawika, Z. Alyafeai, A. Chaffin, A. Stiegler, T. Le Scao, A. Raja, 等, 《多任务提示训练实现零样本任务泛化》, 载于：ICLR 2022——第十届国际学习表征会议, 2022。[178] P. Törnberg, D. Valeeva, J. Uitermark, C. Bail, 《利用大语言模型模拟社交媒体, 以评估替代性新闻推送算法》, arXiv预印本 arXiv:2310.05984（2023年）。[179] Z. Yang, Z. Zhang, Z. Zheng, Y. Jiang, Z. Gan, Z. Wang, Z. Ling, M. Ma, B. Dong, P. Gupta, 等, 《OASIS：大规模开放式智能体社交互动模拟》（2025年）。arXiv:2411.11581。

[180] X. Liu, A. Nourbakhsh, Q. Li, R. Fang, S. Shah，在 Twitter 上实时辟谣，载于：第24届 ACM 信息与知识管理国际会议论文集，2015年，第1867–1870页。[181] J. Ma, W. Gao, P. Mitra, S. Kwon, B. J. Jansen, K.-F. Wong, M. Cha，使用循环神经网络从微博中检测谣言（2016）。[182] A. Zubiaga, M. Liakata, R. Procter，在社交媒体突发新闻期间学习报道动态以检测谣言，arXiv 预印本 arXiv:1610.07363（2016）。[183] L. Derczynski, K. Bontcheva, M. Liakata, R. Procter, G. W. S. Hoi, A. Zubiaga，SemEval-2017任务8：RumourEval——判定谣言真实性及其支持情况，载于：第11届语义评估国际研讨会论文集（SemEval-2017），2017年，第69–76页。[184] G. Gorrell, E. Kochkina, M. Liakata, A. Aker, A. Zubiaga, K. Bontcheva, L. Derczynski，SemEval-2019任务7：RumourEval——判定谣言真实性及其支持情况，载于：J. May, E. Shutova, A. Herbelot, X. Zhu, M. Apidianaki, S. M. Mohammad（编），第13届语义评估国际研讨会论文集，计算语言学协会，明尼阿波利斯，明尼苏达州，美国，2019年，第845–854页。doi:10.18653/ v1/S19-2147。URL https://aclanthology.org/S19-2147/[185] X. Hu, Z. Guo, J. Chen, L. Wen, P. S. Yu，MR2：用于社交媒体多模态检索增强谣言检测的基准数据集，载于：第46届 ACM 信息检索研究与发展国际会议论文集，2023年，第2901–2912页。[186] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, H. Liu，FakeNewsNet：一个包含新闻内容、社会情境及时空信息、用于研究社交媒体假新闻的数据存储库，《大数据》8（3）（2020）171–188。[187] Q. Nan, J. Cao, Y. Zhu, Y. Wang, J. Li，Mdfend：多领域假新闻检测，载于：第30届 ACM 信息与知识管理国际会议论文集，2021年，第3343–3347页。[188] E. Min, Y. Rong, Y. Bian, T. Xu, P. Zhao, J. Huang, S. Ananiadou，分而治之：用于社交媒体假新闻检测的帖子—用户交互网络，载于：2022年ACM Web会议论文集，2022年，第1148–1158页。[189] P. Qi, Y. Bu, J. Cao, W. Ji, R. Shui, J. Xiao, D. Wang, T.-S. Chua，FakeSV：一个具有丰富社会情境、用于短视频平台假新闻检测的多模态基准数据集，载于：AAAI人工智能会议论文集，第37卷，2023年，第14444–14452页。[190] B. Hu, Q. Sheng, J. Cao, Y. Zhu, D. Wang, Z. Wang, Z. Jin，借鉴过去，面向未来：预测用于假新闻检测的时间趋势，载于：第61届计算语言学协会年会论文集（第5卷：产业专题），2023年，第116–125页。[191] S. Dadkhah, X. Zhang, A. G. Weismann, A. Firouzi, A. A. Ghorbani，规模最大的真实/虚假内容社交媒体真实标注数据集：TruthSeeker，《IEEE计算社会系统汇刊》11（3）（2023）3376–3390。[192] Y. Li, H. He, J. Bai, D. Wen，MCFEND：用于中文假新闻检测的多源基准数据集，载于：2024年ACM Web会议论文集，2024年，第4018–4027页。[193] D. Y. Zhang, L. Shang, B. Geng, S. Lai, K. Li, H. Zhu, M. T. Amin, D. Wang，FauxBuster：利用社交媒体评论进行检测的无内容虚假摄影检测器，载于：2018年IEEE国际大数据会议（大数据），IEEE，2018年，第891–900页。[194] D. S. Nielsen, R. McConville，MuMiN：一个大规模多语言、多模态、经事实核查的错误信息社交网络数据集，载于：第45届 ACM 信息检索研究与发展国际会议论文集，2022年，第3141–3153页。[195] Q. Cao, H. Shen, K. Cen, W. Ouyang, X. Cheng，DeepHawkes：弥合信息级联预测与理解之间的鸿沟，载于：2017年 ACM 信息与知识管理会议论文集，CIKM '17，计算机协会，纽约，美国，2017年，第1149–1158页。doi:10.1145/3132847.3132973。URL https://doi.org/10.1145/3132847.3132973[196] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, A. Scala，新冠疫情社交媒体信息疫情，《科学报告》10（1）（2020）16598。[197] Y. Hu, H.-Y. Huang, A. Chen, X.-L. Mao，Weibo-COV：来自微博的大规模新冠肺炎社交媒体数据集，载于：EMNLP 2020新冠肺炎自然语言处理研讨会（第2部分）论文集，2020年。[198] C. Du, K. Fu, B. Wen, Y. Sun, J. Peng, W. Wei, Y. Gao, S. Wang, C. Zhang, J. Li等，多模态大语言模型中自然涌现的人类式物体概念表征，《自然·机器智能》（2025）1–16。[199] D. Dillion, N. Tandon, Y. Gu, K. Gray，人工智能语言模型能否取代人类参与者？《认知科学趋势》27（7）（2023）597–600。[200] L. P. Argyle, E. C. Busby, N. Fulda, J. R. Gubler, C. Rytting, D. Wingate，由一到众：使用语言模型模拟人类样本，《政治分析》31（3）（2023）337–351。[201] B. Liu, X. Li, J. Zhang, J. Wang, T. He, S. Hong, H. Liu, S. Zhang, K. Song, K. Zhu等，基础智能体的进展与挑战：从脑启发智能到进化、协作与安全系统，arXiv 预印本 arXiv:2504.01990（2025）。

[202] T. Hu, D. Liakopoulos, X. Wei, R. Marculescu, N. J. Yadwadkar，使用大语言模型智能体模拟社交网络中的谣言传播，arXiv 预印本 arXiv:2502.01450（2025）。[203] C. Li, X. Su, H. Han, C. Xue, C. Zheng, C. Fan，量化大语言模型对集体意见动态的影响，arXiv 预印本 arXiv:2308.03313（2023）。[204] T. Ju, Y. Wang, X. Ma, P. Cheng, H. Zhao, Y. Wang, L. Liu, J. Xie, Z. Zhang, G. Liu，基于大语言模型的多智能体社区中被操纵知识的泛滥传播，arXiv 预印本 arXiv:2407.07791（2024）。[205] W. Zheng, X. Tang，使用大语言模型智能体模拟社交网络：信息传播与回音室分析，载于：国际知识与系统科学研讨会，施普林格，2024年，第63–77页。[206] Z. Zhang, J. Lian, C. Ma, Y. Qu, Y. Luo, L. Wang, R. Li, X. Chen, Y. Lin, L. Wu等，TrendSim：使用基于大语言模型的多智能体系统模拟投毒攻击下社交媒体中的热门话题，arXiv 预印本 arXiv:2412.12196（2024）。[207] M. Nasim, S. M. Gilani, A. Qasmi, U. Naseem，使用大语言模型智能体模拟影响力动态，arXiv 预印本 arXiv:2503.08709（2025）。[208] P. Taillandier, J. D. Zucker, A. Grignard, B. Gaudou, N. Q. Huynh, A. Drogoul，将大语言模型整合到基于智能体的社会模拟中：机遇与挑战，arXiv 预印本 arXiv:2507.19364（2025）。[209] B. Sampat, S. Raj，假新闻还是真新闻？了解在社交媒体平台分享假新闻的个体所获得的满足感及其人格特质，《Aslib信息管理期刊》74（5）（2022）840–876。[210] S. Ahmed, M. E. Rasul，社交媒体新闻使用与新冠疫情错误信息参与度：调查研究，《医学互联网研究杂志》24（9）（2022）e38944。[211] A. F. Ashery, L. M. Aiello, A. Baronchelli，大语言模型群体中涌现的社会惯例与集体偏差，《科学进展》11（20）（2025）eadu9368。[212]G. De Marzo, C. Castellano, D. Garcia，人工智能智能体能够实现超越人类尺度的协同，arXiv 预印本 arXiv:2409.02822（2024）。[213] A. Flint, L. M. Aiello, R. Pastor-Satorras, A. Baronchelli，大语言模型多智能体系统中的群体规模效应与集体失调，arXiv 预印本 arXiv:2510.22422（2025）。[214] M. S. Jeon, M. Mendoza, M. Fernández, E. Providel, F. Rodríguez, N. Espina, A. Carvallo, A. Abeliuk，使用生成式基于智能体模型模拟社交媒体对话，《EPJ数据科学》14（1）（2025）79。[215] D. Cirulli, G. Cimini, G. Palermo，大语言模型如何在在线对话中扮演人类：对 Reddit 上2016年美国政治的模拟研究，arXiv 预印本 arXiv:2506.21620（2025）。[216] L. Verginer, G. Vaccario, P. Ronzani，人机交互中的非理性羊群行为依然存在，《科学报告》15（1）（2025）22828。[217] A. Bouleimen, G. De Marzo, T. Kim, N. Pagan, H. Metzler, S. Giordano, D. Garcia，集体图灵测试：大语言模型能够生成逼真的多用户讨论，arXiv 预印本 arXiv:2511.08592（2025）。[218] G. M. Orlando, J. Ye, V. La Gatta, M. Saeedi, V. Moscato, E. Ferrara, L. Luceri，网络化大语言模型智能体中涌现的协同行为：信息行动的战略动态建模，arXiv 预印本 arXiv:2510.25003（2025）。[219] C. Yu, J. Ye, Y. Li, Z. Li, E. Ferrara, X. Hu, Y. Zhao，面向政治学决策的大语言模型大规模模拟，arXiv 预印本 arXiv:2412.15291（2024）。[220] K. R. Bissett, J. Cadena, M. Khan, C. J. Kuhlman，基于智能体的计算流行病学建模，《印度科学院学报》101（3）（2021）303–327。[221] K. Gallagher, I. Bouros, N. Fan, E. Hayman, L. Heirene, P. Lamirande, A. Lemenuel-Diot, B. Lambert, D. Gavaghan, R. Creswell，流行病学基于智能体的建模软件（Epiabm），arXiv 预印本 arXiv:2212.04937（2022）。[222] D. Lakshmi S，生成式人工智能时代的帝国再塑，《传播前沿》10（2025）1604361。[223] B. Jiang, C. Zhao, Z. Tan, H. Liu，捕捉变色龙：检测使用大语言模型生成的演变中的虚假信息，载于：2024年第六届IEEE认知机器智能国际会议（CogMI），IEEE，2024年，第197–206页。[224] S. Farquhar, J. Kossen, L. Kuhn, Y. Gal，使用语义熵检测大语言模型中的幻觉，《自然》630（8017）（2024）625–630。[225] J. S. Park, L. Popowski, C. Cai, M. R. Morris, P. Liang, M. S. Bernstein，社会拟像：为社会计算系统创建填充式原型，载于：第35届 ACM用户界面软件与技术研讨会论文集，2022年，第1–18页。[226] Y. Huang, J. Zhou, W. Dong, W. Li, M. Chi, C. Jiang, W. Wang, S. Deng，解读大语言模型在在线评论中的言语欺骗，《决策支持系统》（2025）114529。[227] D. Barman, Z. Guo, O. Conlan，语言模型的阴暗面：探索大语言模型在多媒体虚假信息生成与传播中的潜力，《机器学习应用》16（2024）100545。

[228] J. Roozenbeek、S. Van Der Linden、B. Goldberg、S. Rathje、S. Lewandowsky，心理接种可提高对社交媒体错误信息的韧性，《科学进展》8 (34) (2022年) eabo6254.[229] M. Omar、S. Soffer、R. Agbareia、N. L. Bragazzi、D. U. Apakama、C. R. Horowitz、A. W. Charney、R. Freeman、B. Kummer、B. S. Glicksberg等，大语言模型医疗决策中的社会人口学偏差，Nature Medicine (2025年) 1–9.[230] B. C. Z. Tan、R. K.-W. Lee，揭示隐性偏见：评估权力不平等社会情境中角色提示的大语言模型回复，载于：2025年美洲分会计算语言学协会会议论文集：人类语言技术（第1卷：长论文），2025年，第1075–1108页.[231] A. Barfar、L. Sommerfeldt，提示驱动的宣传：追踪大语言模型中隐藏的语言策略，《信息处理与管理》63 (2) (2026年) 104403.[232] J. Uyheng、D. Bellutta、K. M. Carley，机器人账号在新冠疫情期间关于种族主义的在线话语中放大并转移仇恨言论，Social Media+ Society 8 (3) (2022年) 20563051221104749.[233] Y. Zhang、K. Sharma、L. Du、Y. Liu，迈向缓解大语言模型时代的错误信息与社交媒体操纵，载于：2024年ACM Web 会议配套论文集，2024年，第1302–1305页.[234] R. Staab、M. Vero、M. Balunovic、M. Vechev，超越记忆化：通过大语言模型推断侵犯隐私，载于：第十二届国际学习表征会议，2024年.[235] E. Hitz、M. Feng、R. Tanase、R. Algesheimer、M. S. Mariani，人工代理在社会传染中的放大效应，arXiv预印本 arXiv:2502.21037 (2025年).[236] I. Rahwan、M. Cebrian、N. Obradovich、J. Bongard、J.-F. Bonnefon、C. Breazeal、J. W. Crandall、N. A. Christakis、I. D. Couzin、M. O. Jackson等，机器行为，《自然》568 (7753) (2019年) 477–486.[237] D. Bzdok、A. Thieme、O. Levkovskyy、P. Wren、T. Ray、S. Reddy，大语言模型为神经科学与生物医学带来的数据科学机遇，《神经元》112 (5) (2024年) 698–717.[238] Z. Zhang、H. Miao、Y. Liang、Y. Zhao、X. Han、P. Jia、B. Yang、C. S. Jensen，以网络为中心的人类移动分析：大语言模型时代的方法、应用与未来方向，载于：2025年ACM网络会议配套论文集，2025年，第81–84页.[239] P. C. Pinto、P. Thiran、M. Vetterli，在大规模网络中定位扩散源，《物理评论快报》109 (6) (2012年) 068702.[240] K. Zhu、Z. Chen、L. Ying，一网打尽：在部分观测的网络中定位多个扩散源，载于：AAAI人工智能会议论文集，第31卷，2017年.[241] M. Scheffer、J. Bascompte、W. A. Brock、V. Brovkin、S. R. Carpenter、V. Dakos、H. Held、E. H. Van Nes、M. Rietkerk、G. Sugihara，临界转变的早期预警信号，《自然》461 (7260) (2009年) 53–59.[242] R. L. Berkelman、R. T. Bryan、M. T. Osterholm、J. W. LeDuc、J. M. Hughes，传染病监测：正在崩塌的基础，《科学》264 (5157) (1994年) 368–370.[243] M. M. Mello、C. J. Wang，数字疾病监测的伦理与治理，《科学》368 (6494) (2020年) 951–954.[244] J. S. Brownstein、B. Rader、C. M. Astley、H. Tian，人工智能在传染病监测中的进展，《新英格兰医学杂志》388 (17) (2023年) 1597–1607.[245] A. Bodaghi、K. A. Schmitt、P. Watine、B. C. Fung，关于检测、核实与缓解在线错误信息的文献综述，IEEE计算社会系统汇刊11 (4) (2023年) 5119–5145.[246] M. S. Jahan 、 M. Oussalah，利用自然语言处理自动检测仇恨言论的系统综述，《神经计算》546 (2023年) 126232.[247] S. Jiang、W. Tang、X. Chen、R. Tang、H. Wang、W. Wang，ReZG：面向仇恨言论的检索增强零样本反叙事生成，《神经计算》620 (2025年) 129140.[248] C. Yuan、Q. Ma、W. Zhou、J. Han、S. Hu，基于弱监督学习利用新闻、出版商和用户的可信度实现假新闻早期检测，载于：第28届国际计算语言学会议论文集，2020年，第5444–5454页.[249] K. A. Qureshi、R. A. S. Malick、M. Sabih、H. Cherifi，社交媒体上的欺骗检测：基于来源的视角，Knowledge-Based Systems 256 (2022年) 109649.[250] L. Sun、Y. Rao、L. Wu、X. Zhang、Y. Lan、A. Nazir，从传播过程对抗虚假信息：综述，《ACM计算综述》55 (10) (2023年) 1–38.[251] Z. Chen、J. Freire，从实时社交媒体信息流中主动发现假新闻域名，载于：2020年Web会议配套论文集，2020年，第584–592页.[252] H. A. Simon等，人类行为不变量，《心理学年度评论》41 (1) (1990年) 1–20.

[253] E. R. Smith, M. A. Zarate, 基于范例的社会判断模型，《心理学评论》99 (1) (1992) 3.

[254] M. Galesic、W. Bruine de Bruin、J. Dalege、S. L. Feld、F. Kreuter、H. Olsson、D. Prelec、D. L. Stein、T. van Der Does，人类社会感知是计算社会科学尚未开发的资源，《自然》595 (7866) (2021年) 214–222.[255] R. B. Cialdini、N. J. Goldstein，社会影响：顺从与从众，Annu. Rev. Psychol. 55 (1) (2004年) 591–621.[256] X. Lan、C. Gao、D. Jin、Y. Li，利用融入协作角色的基于大语言模型的智能体进行立场检测，载于：国际AAAI网络与社交媒体会议论文集，第18卷，2024年，第891–903页.[257] Z. Liu、K. Yang、Q. Xie、T. Zhang、S. Ananiadou，EmoLLMs：用于全面情感分析的一系列情感大语言模型与标注工具，载于：第30届ACM SIGKDD知识发现与数据挖掘会议论文集，2024年，第5487–5496页.[258] Y. Zhang、M. Wang、Y. Wu、P. Tiwari、Q. Li、B. Wang、J. Qin，DialogueLLM：面向对话中情绪识别的语境与情绪知识调优大语言模型，Neural Networks (2025年) 107901.[259] A. Browne、D. Butts、E. Jaramillo-Rodriguez、N. Parikh、G. Fairchild、Z. Needell、C. Poliziani、T. Wenzel、T. C. Germann、S. Del Valle，评估不同社区结构接触网络中用于疫情暴发早期检测的疾病监测策略，《社交网络》79 (2024年) 122–132.[260] J. L. Herrera、R. Srinivasan、J. S. Brownstein、A. P. Galvani、L. A. Meyers，复杂社会网络上的疾病监测，《PLoS计算生物学》12 (7) (2016) e1004928.[261] Y. Si、Z. Zhao、R. Chen、H. Zhong、T. Liu、M. Wang、X. Song、W. Li、B. Ying，中国西南地区疑似COVID-19患者中常见呼吸道病毒的流行病学监测，《BMC传染病》20 (1) (2020年) 688.[262] S. M. Lemon、M. A. Hamburg、P. F. Sparling、E. R. Choffnes、A. Mack，全球传染病监测与检测：评估挑战——寻求解决方案，研讨会摘要 (2007年).[263] M. Yuan、N. Boston-Fisher、Y. Luo、A. Verma、D. L. Buckeridge，公共卫生监测中所用异常检测算法的系统综述，《生物医学信息学杂志》94 (2019年) 103181.[264] J. Murray、A. L. Cohen，传染病监测，《国际公共卫生百科全书》(2016) 222.[265] W. H. Organization，面向流行病和大流行病的未来监测：2023年视角，世界卫生组织，2023年.[266] E. C. Clark、S. Neumann、S. Hopkins、A. Kostopoulos、L. Hagerman、M. Dobbins，新冠疫情导致的公共卫生监测方法变化：范围综述，《JMIR公共卫生与监测》10 (1) (2024年) e49185.[267] W. Yang，传染病疫情暴发早期预警：理论与实践，学术出版社，2017年.[268] S. Declich、A. O. Carter，公共卫生监测：历史起源、方法与评估，《世界卫生组织公报》72 (2) (1994年) 285.[269] P. Birello、M. Re Fiorentin、B. Wang、V. Colizza、E. Valdano，由疫情监测估计的再生数在空间结构化种群中可能存在偏倚，《自然·物理》20 (7) (2024年) 1204–1210.[270] S. W. Lo、D. Jamrozy，基因组学与流行病学监测，《自然综述：微生物学》18 (9) (2020年) 478–478.[271] D. J. Ecker、R. Sampath、L. B. Blyn、M. W. Eshoo、C. Ivy、J. A. Ecker、B. Libby、V. Samant、K. A. Sannes-Lowery、R. E. Melton等，用于疫情监测的呼吸道病原体快速鉴定与菌株分型，《美国国家科学院院刊》102 (22) (2005年) 8012–8017.[272] A. Moura、A. Criscuolo、H. Pouseele、M. M. Maury、A. Leclercq、C. Tarr、J. T. Björkman、T. Dallman、A. Reimer、V. Enouf等，单核细胞增生李斯特菌的全基因组群体生物学与流行病学监测，《自然·微生物学》2 (2) (2016) 1–10.[273] W.-H. Hu、H.-M. Sun、Y.-Y. Wei、Y.-T. Hao，全球传染病预警模型：更新综述及新冠疫情的启示，Infectious Disease Modelling (2024年).[274] Y. T. Yang、M. Horneffer、N. DiLisio，挖掘社交媒体与网络搜索以进行疾病检测，《公共卫生研究杂志》2 (1) (2013年) jphr–2013.[275] S. Hashimoto、Y. Murakami、K. Taniguchi、M. Nagai，通过传染病监测在早期阶段检测流行病，《国际流行病学杂志》29 (5) (2000年) 905–910.[276] R. Meckawy、D. Stuckler、A. Mehta、T. Al-Ahdal、B. N. Doebbeling，预警系统在检测传染病疫情暴发中的有效性：系统综述，BMC Public Health 22 (1) (2022年) 2216.[277] E. Southall、T. S. Brett、M. J. Tildesley、L. Dyson，传染病转变的早期预警信号：综述，《英国皇家学会界面杂志》18 (182) (2021年) 20210555.[278] C. R. MacIntyre、X. Chen、M. Kunasekaran、A. Quigley、S. Lim、H. Stone、H.-y. Paik、L. Yao、D. Heslop、W. Wei等，公共卫生中的人工智能：疫情预警系统的潜力，《国际医学研究杂志》51 (3) (2023年) 03000605231159335.

[279] I. Villanueva-Miranda, G. Xiao, Y. Xie, 《人工智能在传染病监测预警系统中的应用：系统综述》, Frontiers in Public Health 13（2025年）1609615。[280] A. Alavi, G. K. Bogu, M. Wang, E. S. Rangan, A. W. Brooks, Q. Wang, E. Higgs, A. Celli, T. Mishra, A. A. Metwally 等, 《利用可穿戴设备数据实时预警新冠和其他压力事件》, Nature Medicine 28（1）（2022年）175–184。[281] S. Unkel, C. P. Farrington, P. H. Garthwaite, C. Robertson, N. Andrews, 《前瞻性检测传染病暴发的统计方法：综述》, Journal of the Royal Statistical Society Series A: Statistics in Society 175（1）（2012年）49–82。[282] C. I. Siettos, L. Russo, 《传染病动力学的数学建模》, Virulence 4（4）（2013年）295–306。[283] M. Kulldorff, R. Heffernan, J. Hartman, R. Assunçao, F. Mostashari, 《用于疾病暴发检测的时空置换扫描统计量》, PLoS Medicine 2（3）（2005年）e59。[284] A. Hohl, E. M. Delmelle, M. R. Desjardins, Y. Lan, 《在美国利用前瞻性时空扫描统计量每日监测新冠》, Spatial and Spatio-temporal Epidemiology 34（2020年）100354。[285] A. Whiteman, M. R. Desjardins, G. A. Eskildsen, J. R. Loaiza, 《调整媒介监测数据后检测巴拿马登革热的时空聚集》, PLoS Neglected Tropical Diseases 13（9）（2019年）e0007266。[286] S. Bhatia, B. Lassmann, E. Cohn, A. N. Desai, M. Carrion, M. U. Kraemer, M. Herringer, J. Brownstein, L. Madoff, A. Cori 等, 《利用数字监测工具近实时绘制传染病传播风险图》, NPJ Digital Medicine 4（1）（2021年）73。[287] Z. Li, F. Meng, B. Wu, D. Kong, M. Geng, X. Qiu, Z. Cao, T. Li, Y. Su, S. Liu, 《回顾传染病预警系统的进展并规划未来》, BMC Public Health 24（1）（2024年）3080。[288] G. S. Williams, E. L. Koua, P. Abdelmalik, F. Kambale, E. Kibangou, J. Nguna, C. Okot, G. Akpan, F. Moussana, J. P. Kimenyi 等, 《评估开放源疫情情报（EIOS）系统在非洲区域早期检测疫情暴发和卫生突发事件中的作用》, BMC Public Health 25（1）（2025年）857。[289] W. H. Organization, 《EIOS技术概述》（2025年）。网址：https://www.who.int/initiatives/eios/eios-technology[290] A. Quigley, M. D. Honeyman, M. H. Stone, R. Dawson, C. R. MacIntyre, 《Epiwatch：人工智能预警系统是疫情监测的宝贵工具》, International Journal of Infectious Diseases 152（2025年）107579。[291] N. Bhadelia, 《EAGER：生物威胁出现、分析与通信网络（BEACON）》, 美国国家科学基金会奖项编号2433726。生物科学部 24（2433726）（2024年）33726。[292] H. I. for Computing, D. Sciences, 《BEACON今日上线，提供人工智能驱动的全球疾病监测》, 网页，发布于4月24日（2025年）。网址：https://www.bu.edu/hic/2025/04/24/beacon-delivers-ai-powered-global-disease

-监测/

[293] M. Carrion，L. C. Madoff，ProMED-mail：对新发传染病进行数字监测的22年历程，《国际卫生》9（3）（2017）177–183。[294] C. R. MacIntyre，S. Lim，A. Quigley，预防下一场大流行病：利用人工智能进行疫情监测和警报，《Cell Reports Medicine》3（12）（2022）。[295] C. C. Freifeld，K. D. Mandl，B. Y. Reis，J. S. Brownstein，HealthMap：通过自动分类和可视化互联网媒体报道进行全球传染病监测，《美国医学信息学协会杂志》15（2）（2008）150–157。[296] J. Ong，X. Liu，J. Rajarethinam，S. Y. Kok，S. Liang，C. S. Tang，A. R. Cook，L. C. Ng，G. Yap，利用随机森林绘制新加坡登革热风险地图，《PLoS Neglected Tropical Diseases》12（6）（2018）e0006587。[297] F. Chen，D. B. Neill，异质社交媒体图中的事件检测与预测非参数扫描统计量，载于：第20届ACM SIGKDD知识发现与数据挖掘国际会议论文集，2014，页1166–1175。[298] V. Lampos，B. Zou，I. J. Cox，利用词嵌入增强特征选择：以流感监测为例，载于：第26届万维网国际会议论文集，2017，页695–704。[299] F. Ding，J. Fu，D. Jiang，M. Hao，G. Lin，绘制埃及伊蚊和白纹伊蚊的空间分布，《热带学报》178（2018）155–162。[300] Z. Jiao，H. Ji，J. Yan，X. Qi，大数据与人工智能在疫情监测和遏制中的应用，《智能医学》3（1）（2023）36–43。[301] A. X. Han，A. Toporowski，J. A. Sacks，M. D. Perkins，S. Briand，M. Van Kerkhove，E. Hannay，S. Carmona，B. Rodriguez，E. Parker，等，SARS-CoV-2诊断检测率决定基因组监测项目的灵敏度，《自然—遗传学》55（1）（2023）26–33。[302] R. Allard，在传染病监测中使用时间序列分析，《世界卫生组织公报》76（4）（1998）327。

[303] L. Tomov，L. Chervenkov，D. G. Miteva，H. Batselova，T. Velikova，时间序列分析在流行病学中的应用：文献综述及我们在新冠疫情期间的经验，《世界临床病例杂志》11（29）（2023）6974。[304] C. Katris，基于时间序列的疫情传播预测统计方法：以希腊COVID-19疫情为例，《专家系统与应用》166（2021）114077。[305] P. Lemey，S. L. Hong，V. Hill，G. Baele，C. Poletto，V. Colizza，Á. O'toole，J. T. McCrone，K. G. Andersen，M. Worobey，等，在SARS-CoV-2贝叶斯系统地理推断中纳入个体旅行史和未采样多样性，《自然·通讯》11（1）（2020）5110。[306] R. Li，S. Pei，B. Chen，Y. Song，T. Zhang，W. Yang，J. Shaman，大量未报告感染助长了新型冠状病毒（SARS-CoV-2）的快速传播，《科学》368（6490）（2020）489–493。[307] X. Wu，J. Gao，M. Bilal，F. Dai，X. Xu，L. Qi，W. Dou，面向物联网疫情监测的基于联邦学习的隐私医疗知识图谱，《专家系统》42（1）（2025）e13372。[308] D. Bang，S. Lim，S. Lee，S. Kim，通过将"因关联而有罪"扩展至多个层，为药物再利用进行生物医学知识图谱学习，《自然·通讯》14（1）（2023）3570。[309] S. Consoli，P. Coletti，P. V. Markov，L. Orfei，I. Biazzo，L. Schuh，N. Stefanovitch，L. Bertolini，M. Ceresa，N. I. Stilianakis，从世界卫生组织疾病暴发新闻中提取的流行病学知识图谱，《科学数据》12（1）（2025）970。[310] F. Stadtmann，E. R. Furevik，A. Rasheed，T. Kvamsdal，物理引导的联邦学习助力数字孪生，《专家系统与应用》258（2024）125169。[311] Y. Song，Y. Wu，S. Wu，D. Li，Q. Wen，S. Qin，F. Gao，面向经典客户端的量子联邦学习框架，《中国科学：物理学、力学、天文学》67（5）（2024）250311。[312] T. Qi，F. Wu，C. Wu，L. He，Y. Huang，X. Xie，面向联邦学习的差分隐私知识迁移，《自然·通讯》14（1）（2023）3785。[313] X. Ma，X. Chang，H. Chen，基于奇异值分解的网络敏感信息差分隐私保护算法，《科学报告》13（1）（2023）6035。[314] K. Cheng，Z. Li，Y. He，Q. Guo，Y. Lu，S. Gu，H. Wu，人工智能在传染病领域的潜在应用：以ChatGPT为例，《生物医学工程年鉴》51（6）（2023）1130–1135。[315] J. Xie，F. Meng，J. Sun，X. Ma，G. Yan，Y. Hu，检测并建模社交媒体上信息的真实渗流与相变，《自然·人类行为》5（9）（2021）1161–1168。[316] M. Granovetter，集体行为的阈值模型，《美国社会学杂志》83（6）（1978）1420–1443。[317] J. Goldenberg，B. Libai，E. Muller，网络中的谈论：从复杂系统视角看口碑传播的底层过程，《营销快报》12（3）（2001）211–223。[318] Q. Cao，H. Shen，J. Gao，B. Wei，X. Cheng，利用耦合图神经网络预测社交平台上的流行度，载于：第13届网络搜索与数据挖掘国际会议论文集，2020，页70–78。[319] T. Wu，L. Chen，X. Xian，Y. Guo，多尺度信息扩散动力学的演化预测，《知识型系统》113（2016）186–198。[320] A. K. Taylor，N. Wen，P.-N. Kung，J. Chen，V. Peng，W. Wang，你的新闻来自哪里？预测社交媒体中的信息传播路径，载于：第46届ACM SIGIR信息检索研究与发展国际会议论文集，2023，页2511–2515。[321] Y. Jin，Y.-C. Lee，K. Sharma，M. Ye，K. Sikka，A. Divakaran，S. Kumar，预测在线社群间的信息传播路径，载于：第29届ACM SIGKDD知识发现与数据挖掘会议论文集，2023，页1044–1056。[322] L. R. Goldberg，"人格描述"的另一种方式：大五人格因素结构，载于：《人格与人格障碍》，劳特利奇，2013，页34–47。[323] 欧洲议会和欧盟理事会关于应对恐怖主义内容在网上传播的2021年4月29日欧盟条例2021/784，《欧盟官方公报》L 172，页79–109，第3(3)条（2021）。网址：https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32021R0784[324] 世界卫生组织，信息疫情管理，访问日期：2025年10月5日（2024）。网址：https://www.who.int/health-topics/infodemic[325] S. Lewandowsky，J. Cook，U. Ecker，D. Albarracin，M. A. Amazeen，P. Kendou，D. Lombardi，E. Newman，G. Pennycook，E. Porter，等，《2020年辟谣手册》，2020。[326] A. M. Guess，M. Lerner，B. Lyons，J. M. Montgomery，B. Nyhan，J. Reifler，N. Sircar，一项数字媒体素养干预提升了美国和印度受众区分主流新闻与虚假新闻的能力，《美国国家科学院院刊》117（27）（2020）15536–15545。[327] G. Pennycook，D. G. Rand，准确性提示是一种可重复且可推广的方法，有助于减少错误信息的传播，《自然·通讯》13（1）（2022）2333。

[328] T. K. Tsang, Q. Du, B. J. Cowling, C. Viboud，在季节性不规律的情况下采用自适应权重集成方法预测流感活动，《自然·通讯》15 (1) (2024) 8625。[329] J. B. Bak-Coleman, M. Alfano, W. Barfuss, C. T. Bergstrom, M. A. Centeno, I. D. Couzin, J. F. Donges, M. Galesic, A. S. Gersick, J. Jacquet 等，《全球集体行为的管理》，《美国国家科学院院刊》118 (27) (2021) e2025764118。[330] A. Cho，人工智能系统旨在嗅探冠状病毒疫情暴发，《科学》368 (6493) (2020) 810–811。doi:10.1126/ science.368.6493.810。[331] S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch，预测后疫情时期SARS-CoV-2的传播动力学，《科学》368 (6493) (2020) 860–868。[332] R. M. Anderson, R. M. May，传染病的种群生物学：第一部分，《自然》280 (5721) (1979) 361–367。[333] X. Li, H. Liu, L. Gao, S. P. Sherchan, T. Zhou, S. J. Khan, M. C. Van Loosdrecht, Q. Wang，基于废水的流行病学预测了COVID-19导致的美国150多个县每周新增住院人数，《自然·通讯》14 (1) (2023) 4548。[334] M. Tizzoni, P. Bajardi, C. Poletto, J. J. Ramasco, D. Balcan, B. Gonçalves, N. Perra, V. Colizza, A. Vespignani，全球流行病传播的实时数值预测：2009年A/H1N1pdm案例研究，BMC Medicine 10 (1) (2012) 165。[335] C. Liu, X.-X. Zhan, Z.-K. Zhang, G.-Q. Sun, P. M. Hui，事件如何决定传播模式：社交网络中通过内外部影响进行的信息传播，《新物理学杂志》17 (11) (2015) 113045。[336] M. R. Nepomuceno, E. Acosta, D. Alburez-Gutierrez, J. M. Aburto, A. Gagnon, C. M. Turra，除了人口年龄结构之外，健康及其他人口因素也有助于了解COVID-19负担，《美国国家科学院院刊》117 (25) (2020) 13881–13883。[337] A. Telenti, A. Arvin, L. Corey, D. Corti, M. S. Diamond, A. García-Sastre, R. F. Garry, E. C. Holmes, P. S. Pang, H. W. Virgin，大流行病之后：对COVID-19未来发展轨迹的展望，《自然》596 (7873) (2021) 495–504。[338] D. B. Searls，基因的语言，《自然》420 (6912) (2002) 211–217。[339] K. Ruggeri, F. Stock, S. A. Haslam, V. Capraro, P. Boggio, N. Ellemers, A. Cichocka, K. M. Douglas, D. G. Rand, S. Van der Linden 等，新冠期间行为科学为政策提供的证据综述，《自然》625 (7993) (2024) 134–147。[340] J. P. Ioannidis, S. Cripps, M. A. Tanner，COVID-19预测未能奏效，《国际预测学杂志》38 (2) (2022) 423–438。[341] H. McClymont, S. B. Lambert, I. Barr, S. Vardoulakis, H. Bambrick, W. Hu，基于互联网的监测系统与传染病预测：对过去10年的最新综述及新冠疫情的经验教训，《流行病学与全球健康杂志》14 (3) (2024) 645–657。[342] E. Ferrara，生成式人工智能反人类：生成式人工智能和大语言模型的恶意应用，《计算社会科学杂志》7 (1) (2024) 549–569。[343] Y. Chang, X. Wang, J. Wang, Y. Wu, L. Yang, K. Zhu, H. Chen, X. Yi, C. Wang, Y. Wang 等，大语言模型评估综述，ACM智能系统与技术汇刊15 (3) (2024) 1–45。[344] J. Chen, H. Lin, X. Han, L. Sun，大语言模型在检索增强生成中的基准测试，载于：AAAI人工智能会议论文集，第38卷，2024年，第17754–17762页。[345] Q. Chen, Y. Hu, X. Peng, Q. Xie, Q. Jin, A. Gilson, M. B. Singer, X. Ai, P.-T. Lai, Z. Wang 等，面向生物医学自然语言处理应用的大语言模型基准测试与建议，《自然·通讯》16 (1) (2025) 3280。[346] L. Zhong, Z. Wang，大语言模型能取代Stack Overflow吗？关于大语言模型代码生成鲁棒性与可靠性的研究，载于：AAAI人工智能会议论文集，第38卷，2024年，第21841–21849页。[347] X. Qi, K. Huang, A. Panda, P. Henderson, M. Wang, P. Mittal，视觉对抗样本可使经过对齐的大语言模型越狱，载于：AAAI人工智能会议论文集，第38卷，2024年，第21527–21536页。[348] S. Jiang, X. Chen, K. Xu, L. Chen, H. Ren, R. Tang，分解、合成与攻击：一种用于使大语言模型越狱的多指令融合方法，《IEEE物联网杂志》（2025）。[349] S. Jiang, X. Chen, R. Tang，利用带有隐蔽攻击的组合式指令欺骗大语言模型，《ACM自主与自适应系统汇刊》（2025）。[350] L. Huang, W. Yu, W. Ma, W. Zhong, Z. Feng, H. Wang, Q. Chen, W. Peng, X. Feng, B. Qin 等，大语言模型幻觉综述：原理、分类体系、挑战与开放问题，《ACM信息系统汇刊》43 (2) (2025) 1–55。