

THIS EDITION

五个方向的当日进展

收益不确定性下城市服务设施空间布局优化：一种贝叶斯引导的分层强化学习框架

研究日报 · 聚焦地理大模型驱动的自主系统与韧性城市实践。

Volume 19, Issue 2, December 2026 .

编者按：本期头版聚焦‘地理大模型与地理智能体’这一主导趋势，其影响力已贯穿交通优化、设施布局、城市感知与多智能体协同等方向。所选条目体现从静态表征向动态决策、从单点工具向可组合智能体底座的关键跃迁。

TREND OVERVIEW

趋势综述：地理智能体崛起：从空间建模到城市决策闭环。

近期研究重心正从单纯提升GeoFM在遥感影像任务上的准确率，转向增强其对人类活动、功能结构与行为语义的表征能力，并探索其作为可组合、可自主调度的地理智能体（Geo-Agent）底座的潜力。

近期研究重心从两两模态融合转向统一、可扩展的多模态联合建模，强调以场景内容为锚点的跨模态生成与理解；方法上日益依赖大模型架构，并注重对异构传感器物理特性、空间粒度与语义鸿沟的显式建模。

近期研究重心从静态路径规划转向基于实时轨迹反馈的自适应、去中心化交通控制与协同决策；方法上强调神经网络替代表格型学习、移动轮廓（movement profile）替代离散动作、多智能体协同替代单点优化。

DIRECTION PULSE

1 地理大模型与地理智能体

近期研究重心正从单纯提升GeoFM在遥感影像任务上的准确率，转向增强其对人类活动、功能结构与行为语义的表征能力，并探索其作为可组合、可自主调度的地理智能体（Geo-Agent）底座的潜力。

2 多源多模态地理数据

近期研究重心从两两模态融合转向统一、可扩展的多模态联合建模，强调以场景内容为锚点的跨模态生成与理解；方法上日益依赖大模型架构，并注重对异构传感器物理特性、空间粒度与语义鸿沟的显式建模。

3 轨迹数据与城市交通研究

近期研究重心从静态路径规划转向基于实时轨迹反馈的自适应、去中心化交通控制与协同决策；方法上强调神经网络替代表格型学习、移动轮廓（movement profile）替代离散动作、多智能体协同替代单点优化。

4 复杂网络、韧性城市与地理模拟

近期研究聚焦于将复杂网络结构建模与地理模拟技术耦合，以刻画城市系统在气候冲击与人为扰动下的动态韧性响应；方法重心正从静态网络表征转向多尺度、多协议、可干预的高保真仿真平台构建。

5 城市感知、街景感知与空间优化

近期研究重心从静态街景图像的单帧感知建模，转向关注感知的动态性、主体性与行为闭环——尤其强调局部视觉输入如何支撑真实尺度城市中的空间定位、导航与鲁棒决策。方法上正融合MLLM、三维地理仿真与人类感知机制（如注视）以弥合感知与行动鸿沟。

HIGHLIGHTS

- 地理智能体正成为连接空间数据、城市服务与实时决策的核心架构。
- 多源交通数据集支撑从燃油经济性到安全缺陷识别的双重治理路径。
- 贝叶斯引导的强化学习框架为不确定环境下的设施布局提供新范式。
- 面向街景与轨迹的鲁棒感知方法开始融入人类行为闭环建模。

近期研究重心正从单纯提升GeoFM在遥感影像任务上的准确率，转向增强其对人类活动、功能结构与行为语义的表征能力，并探索其作为可组合、可自主调度的地理智能体（Geo-Agent）底座的潜力。

近30天 153 近7天 33 来源 64 论文 1266

趋势信号

- MoRAX与BEACON等新方法显式引入人类移动轨迹和POI行为数据，以弥补EO数据在功能表征上的缺失
- 多篇论文指出GeoFM在分布偏移下校准性差、鲁棒性未显著优于ImageNet预训练模型，引发对评估范式的反思
- 行星预测引擎（PPE）提出端到端自主工作流，将GeoFM嵌入与智能数据检索、模型搜索耦合，体现‘地理智能体’雏形
- Clay等GeoFM开始被用作辅助上下文而非主干编码器，体现向模块化、混合架构演进 的实践信号

核心观点

- 地理空间基础模型（GeoFM）当前核心局限在于过度依赖地球观测（EO）数据，导致对人类活动、城市功能与社会语义的表征薄弱
- 仅依赖准确率（如平均排名）评估GeoFM存在严重缺陷；校准性（calibration）与分布偏移鲁棒性是影响实际部署可信度的关键维度
- 增强GeoFM表征能力的有效路径是融合多源异构信号——特别是人类移动性（Mobility）、POI语义与行为时序数据——而非仅扩展EO模态
- GeoFM正从静态表征工具转向智能体系统组件：其价值不仅在于嵌入质量，更在于能否支撑自主数据合成、任务适配与跨域零样本迁移

RESEARCH IDEA

地理基础模型跨场景迁移的关键瓶颈

地理基础模型迁移到新城市、新尺度或新数据源时，关键瓶颈通常来自空间先验不足、标注差异和工具调用能力不稳定。

为什么现在值得做：行星预测引擎（PPE）已实现多源地理数据的自主检索与融合，可支撑构建覆盖非发达国家的增强型街景-POI-地形三模态对齐数据集；该结果将直接服务于联合国可持续发展目标（SDG）11.3.1（包容性城市规划）与灾害响应中的快速位置验证需求。

关键难点

- 需定义区域级地理定位的可验证ground truth：不能依赖GPS坐标，而需基于行政边界、功能区划与本地地名知识图谱联合校准
- 街景图像中发展中国家典型干扰项（如临时摊位、非标路牌、植被遮挡）缺乏统一标注协议
- GeoAgent的导航策略模块与定位模块耦合紧密，难以分离评估是感知偏差还是路径规划导致的定位漂移

建议切入

- 基于BEACON框架中POI文本-行为-物理嵌入对齐方法，构建欠覆盖城市街景图像的三模态弱监督标签：用AlphaEarth嵌入作空间锚点，POI语义作功能约束，小时级访问量作行为先验
- 在GeoAgent导航策略中插入空间不确定性门控机制，强制智能体在进入低POI密度或高建筑年代混杂性区域时触发多步回溯观测
- 采用MoRAX提供的跨城市功能连通性先验，重加权GeoAgent的区域级候选地理单元得分，替代原始VLM的全局相似度排序

REPRESENTATIVE ITEMS

ISPRS JOURNAL OF PHOTOGRAMMETRY AND REMOTE SENSING

衡量GeoAI与地理空间基础模型的空间显式性：系统性综述与未来路线图

出版日期：2026年10月；来源：《ISPRS摄影测量与遥感杂志》（ISPRS Journal of Photogrammetry and Remote Sensing），第240卷；作者：周 Bing、黄 Xiao、王 Siqin、麦 Gengchen、李 Diya、王 Zhangyu、吴 Qiusheng、宁 Huan、王 Junbo、吴 Qifan、贾 Yuhao、王 Ruomei、吕 Shaokun、唐 Bolong、黄 Zixin。

ARXIV

MoRAX：基于移动性的地理空间基础模型表征增强方法

地理空间基础模型（GFMs）正成为学习语义丰富且地理一致的视觉与物理表征的一种强大范式。然而，其对地球观测（EO）数据的依赖导致人类活动相关信息在很大程度上未被充分表征。人类移动数据揭示了区域间功能与关系结构，而此类信息在EO数据中缺失；但该类数据通常仅覆盖单一城市，因而难以用于可迁移的城市表征学习。

ARXIV

超越准确率：评估地理空间基础模型（GeoFM）s

地理空间基础模型（GeoFM）s通常仅依据其在标准基准条件下的准确率（通过平均排名）进行排序与选择。本文指出，该评估范式过于狭窄：面向关键地球观测（EO）任务的实际部署，亟需引入额外分析维度，尤其是校准性（calibration），即模型置信度与其预测正确性之间的一致性。我们在16个冻结编码器、4个分类数据集与5个分割数据集、以及两条正交压力轴上开展实验，结果表明：所有编码器均随扰动强度增加而性能下降，且其相对排名亦随之变化。

ARXIV

BEACON：通过三模态对比学习实现AlphaEarth嵌入

诸如AlphaEarth Foundation等地理空间基础模型可生成紧凑且全球一致的地球表面表征，能有效迁移到多种下游任务。然而，由于这些模型主要基于地球观测影像进行训练，其嵌入主要捕获物理与光谱特征，而对人类活动与城市功能的编码则较弱。

近期研究重心从两两模态融合转向统一、可扩展的多模态联合建模，强调以场景内容为锚点的跨模态生成与理解；方法上日益依赖大模型架构，并注重对异构传感器物理特性、空间粒度与语义鸿沟的显式建模。

近30天 239

近7天 59

来源 58

论文 1936

趋势信号	核心观点
- MetaEarth-MM 提出场景中心联合建模范式，解耦潜在场景表征与模态生成，支持五种遥感模态间任意翻译 - Earth-OneVision 构建统一 RS-MLLM 框架，覆盖六类传感器模态与九类任务，并设计 FGVLA、SLIS、PCMA 三项专用机制 - TSMNet 引入双分支文本编码器，分别提取物体级标签与场景级语义，通过文本引导视觉语义融合模块实现开放词汇分割 - SA-GEM 提出尺度自适应与地理空间证据调制的令牌剪枝方法，强调令牌质量优于数量，需任务对齐的粒度自适应	- 多模态遥感数据的内在一致性源于底层地理场景，而非表面像素映射，因此场景中心建模比两两模态翻译更具泛化性与可扩展性 - 文本模态（如物体级标签、场景级描述）是弥合遥感视觉语义鸿沟的关键知识源，其引入显著提升开放词汇理解与人类可解释性 - 传感器模态间存在本质差异（成像物理、视角、分辨率），统一建模必须显式处理这些异构性，而非简单特征拼接或对齐 - 高分辨率不等于高性能：在遥感大模型中，地理空间关键证据稀疏且易被冗余令牌稀释，令牌质量与证据保留优先于数量扩展

RESEARCH IDEA

SAM3在遥感零样本分割中因视觉提示与文本提示耦合解码失效

SAM3的多模态解码器在遥感零样本分割任务中，因视觉提示与文本提示共享同一解码路径导致跨模态干扰，无法独立响应语义查询或几何提示

为什么现在值得做：灾害应急与快速制图场景亟需免训练的开放词汇分割能力；Earth-OneVision与TSMNet已验证文本-视觉联合建模的可行性，为解耦式解码器设计提供架构参照；SA-GEM证明令牌级任务对齐可行，支撑提示粒度可控的干预接口。

关键难点	建议切入
- 需定位SAM3解码器中视觉/文本token混合嵌入的具体层与位置 - 需构造隔离视觉提示或文本提示的反事实推理测试协议 - 需定义遥感特异性提示冲突指标（如俯视几何敏感度下降率	- 基于SAM3原始权重冻结主干，插入可学习的模态门控模块以分离视觉/文本token流 - 在五类遥感典型场景（城市密集区、农田斑块、滑坡体、港口设施、林火迹地）上构建提示冲突基准子集 - 采用TSMNet的双分支文本编码器输出作为文本提示源，避免重训语言编码器

REPRESENTATIVE ITEMS

ARXIV

MetaEarth-MM

多模态遥感图像对地球观测至关重要，但在实际应用中，完整的配对观测往往稀缺。现有生成方法通常通过孤立的两两模态翻译来应对该问题，但随着模态数量与生成任务种类的增加，其通用性与可扩展性仍显不足。本文提出一种面向多模态遥感影像的生成式基础模型 MetaEarth-MM，支持在统一框架下实现五种模态间的配对联合生成及任意模态到任意模态的翻译。

ARXIV

融合物体级标签与场景级语义特征的开放词汇语义分割网络

多模态遥感图像的语义分割在土地利用/土地覆盖（LULC）制图、环境监测及精准地球观测中发挥着关键作用。当前多模态方法主要集中于融合互补的视觉模态，却忽视了非视觉文本数据这一富含知识的信息源——文本可有效弥合视觉模式与现实世界概念之间的语义鸿沟。为解决该局限，我们提出TSMNet：一种文本监督的多模态开放词汇语义分割网络，通过协同整合文本监督与视觉表征实现开放词汇语义分割。

ARXIV

Earth-OneVision

遥感多模态大语言模型（RS-MLLMs）支持对地球观测影像的自然语言理解与空间推理。然而，现有模型仅支持有限的传感器类型与任务，导致对地球的观测呈现碎片化，并使跨模态地球科学知识在很大程度上未被利用。本工作提出 Earth-OneVision，一个参数量为20亿的 RS-MLLM，其在单一自回归框架内统一了六类传感器模态（即光学、合成孔径雷达 SAR、红外、多光谱、时序、视频）以及涵盖九类任务的跨传感器融合能力。

INTERNATIONAL JOURNAL OF APPLIED EARTH OBSERVATION AND GEOINFORMATION

面向遥感语义分割的视觉基础模型参数高效多模态适配

出版日期：2026年9月；来源：《国际应用地球观测与地理信息学杂志》（International Journal of Applied Earth Observation and Geoinformation），第153卷；作者：李浩成、郑珏鹏、苗爽熙、卢瑞博、蔡国胜、傅昊桓、黄建熙。

近期研究重心从静态路径规划转向基于实时轨迹反馈的自适应、去中心化交通控制与协同决策；方法上强调神经网络替代表格型学习、移动轮廓（movement profile）替代离散动作、多智能体协同替代单点优化。

近30天 480	近7天 127	来源 70	论文 3189
----------	---------	-------	---------

趋势信号

- 神经双Q路由以共享状态－动作价值网络替代目标索引的表格型Q值学习，解决稀疏访问下信息无法迁移的问题
- VeloCity框架将时空移动轮廓（movement profile）规划直接下放至CAV个体，通过本地预约机制实现去中心化冲突消解
- MINT-V2X数据集首次耦合SUMO交通仿真与OMNeT++/Simu5G网络仿真，提供车辆轨迹与无线信道参数的同步统一视图
- TrafficSci将交通规律发现建模为可审计的AI代理 workflow，实现从轨迹数据中自主归纳并验证跨城市一致的驾驶行为时间记忆尺度

核心观点

- 轨迹数据已超越传统OD分析或可视化用途，成为建模动态交通交互、运动学约束与通信耦合效应的核心输入
- 去中心化不等于无协调：VeloCity和RL信号控制均依赖轻量级本地－协调器交互机制，在降低通信开销的同时保障系统级一致性
- 仿真驱动已成为该方向方法验证的刚性前提——所有六篇论文均明确使用SUMO、POLARIS或自建高保真仿真环境进行评估
- 多智能体范式正从‘小规模预定义场景’向‘任意拓扑城市区域’扩展，但其可扩展性仍需依赖结构化时空抽象（如移动轮廓、预约槽位）而非原始轨迹序列

RESEARCH IDEA

神经双Q路由在城市路网中失效于路口冲突消解机制缺失

神经双Q路由算法在OHT系统中通过剩余回报回归学习状态－动作价值，但在城市道路交叉口场景下因未建模车辆运动学约束与多向冲突消解逻辑，导致其生成的移动轮廓在物理可执行性上不可靠

为什么现在值得做：城市CAV部署正从封闭厂区向开放路网扩展，交通管理部门需可验证的物理安全路由输出；MINT-V2X提供了首个同步轨迹－网络参数的城市级仿真数据源，支持对运动学约束嵌入效果的定量评估。

关键难点

- 需将OHT系统的离散节点图结构映射为城市路口的连续转向运动学模型
- 剩余回报回归目标函数未定义冲突惩罚项，无法直接适配路口多向通行博弈
- 缺乏公开的城市级带运动学标注的冲突事件轨迹真值集用于监督微调

建议切入

- 基于SUMO导出路口转向轨迹的微分代数方程约束集，构建运动学可行域掩码
- 在神经双Q网络输出层后插入轻量级冲突图解码器，以MINT-V2X中CQI-SINR时序突变点为冲突代理标签
- 设计两阶段训练：先用MINT-V2X中15个RSU覆盖区域的987万条同步数据预训练冲突感知表征，再在VeloCity仿真环境中微调路由策略

REPRESENTATIVE ITEMS

ARXIV

基于轨迹初始化的神经双Q路由算法用于大规模天车搬运系统
大规模工业机器人车队共享受限的物理基础设施，导致车辆通行时间受安全间距、路口访问权、下游阻塞及站点争用等因素影响。本文以洁净室半导体制造厂中典型的天花板式物料搬运系统——天车搬运系统（Overhead Hoist Transport, OHT）为研究对象。

ARXIV

基于时空移动轮廓规划的去中心化多智能体城市交通管理
随着现代城市交通拥堵日益加剧，网联自动驾驶车辆（CAVs）已成为下一代智能交通管理的关键使能技术。然而，当前范式存在若干局限，制约了该潜力的充分实现：现有方法通常仅优化局部交互而非系统级效率，通信开销巨大，或缺乏运动学安全执行所必需的确定性保障；此外，当前多智能体方法常局限于小规模预定义场景，难以扩展至大规模复杂城市路网。

ARXIV

基于智能体的流动性与能源政策建模：以2050年芝加哥地区为案例
大都市区正同步推进多项干预措施，以提升交通流动性、可达性及能源效率，亟需集成化工具来理解这些政策如何相互作用，进而影响出行行为、能源消耗及基础设施需求。本文以‘照常情景’（BAU）为基准，评估电气化、货运需求管理、道路收费、停车改革及公共交通扩展等政策组合对2050年芝加哥大都市交通系统的综合影响。我们采用POLARIS——一个基于2019年条件校准的大规模智能体建模框架，在伊利诺伊州东北部七县区域模拟九种政策情景。

ARXIV

基于强化学习的物联网赋能交叉口交通信号控制
城市交通拥堵在以汽车为主导的城市中仍是长期存在的挑战，带来显著的经济与社会成本。交通信号系统正日益作为网络化信息物理系统组件部署于智慧城市基础设施中，分布式感知与边缘智能支撑了自适应交通管理。本文研究强化学习（RL）作为一种边缘智能方法，用于科威特某信号控制城市交叉口的自适应交通信号运行。

近期研究聚焦于将复杂网络结构建模与地理模拟技术耦合，以刻画城市系统在气候冲击与人为扰动下的动态韧性响应；方法重心正从静态网络表征转向多尺度、多协议、可干预的高保真仿真平台构建。

近30天 21 近7天 5 来源 39 论文 260

趋势信号

- 出现轻量级模块化网络靶场（Cyber Range）用于模拟IT/OT融合场景下的城市关键基础设施级攻击链
- 多项研究采用基于智能体的大规模交通-能源耦合建模（如POLARIS框架），显式嵌入政策干预与基础设施响应反馈
- 高原—山地、珠三角等典型地理单元被作为差异化韧性条件下的网络结构演化实证对象
- 地理空间属性与设施属性被联合集成于机器学习模型中，以提升洪涝等灾害下关键基础设施风险的空间精细化识别

核心观点

- 城市韧性不能仅依赖单系统加固，而需建模跨域耦合关系（如交通-能源-通信网络间的协议级依赖）
- 地理异质性（如高原—山地地形、城市群发育阶段）是塑造韧性网络结构与脆弱性分布的根本约束条件
- 高保真地理模拟的价值不仅在于预测，更在于提供可操作的‘干预试验场’——支持绿队评估、政策组合推演与基础设施布局优化
- 韧性评估正从宏观指标（如恢复时间）下沉至网络拓扑特征（如模块度、介数中心性）与协议层行为（如Modbus/TCP中断传播）

RESEARCH IDEA

方法迁移到新场景后的失效边界

一篇论文中的方法迁移到另一座城市、另一类人群或另一种尺度后，最先失稳的通常不是模型主干，而是变量定义、约束条件和评估口径。

为什么现在值得做：城市洪涝频次上升使基础设施物理-信息耦合失效成为应急响应刚需；高分辨率LiDAR水文模型与开源OT协议栈（如pymodbus）已可支撑深度-延迟映射建模，结果可直接服务于应急管理部門的‘断电窗口期’预判。

关键难点

- 需建立淹没深度（cm）到Modbus RTU/TCP超时（ms）的实验室标定关系，现有文献无此物理-协议跨层参数映射
- CR平台YAML引擎不支持动态注入空间栅格变量作为协议行为参数，需修改底层QEMU网络设备仿真模块
- 华盛顿特区实证中未公开设施级离线时间戳序列，需从公开SCADA日志或电力公司 outage reports 中反向推断

建议切入

- 先对两篇代表论文做变量和评价口径对照，确认差异到底来自设定还是方法。
- 再做一个小规模复现或案例迁移，先找出最先失稳的部分。
- 最后根据失稳环节反推真正需要补的数据或约束，而不是直接堆方法。

REPRESENTATIVE ITEMS

SUSTAINABLE CITIES AND SOCIETY

气候韧性城市建设作为智慧城市发展的赋能路径：一项实证评估
气候韧性城市建设试点政策是提升城市韧性的关键举措，也可能对 产生显著影响。

ARXIV

面向智能能源系统的模块化网络靶场平台

现代关键基础设施亟需能够弥合数字漏洞与物理后果之间鸿沟的高保真训练工具。本文提出一种轻量级、模块化的网络靶场（Cyber Range, CR），旨在高保真地模拟信息技术（IT）与运行技术（OT）环境。该平台以FastAPI框架为后端，采用Quick Emulator（QEMU）/基于内核的虚拟机（KVM）虚拟化技术，并通过基于YAML的引擎，在标准硬件上以极低开销自动部署复杂网络。

SUSTAINABLE CITIES AND SOCIETY

基于韧性支撑条件差异的高原—山地城市群城市韧性关联网络
出版日期：2026年10月1日；来源：《可持续城市与社会》（Sustainable Cities and Society），第149卷；作者：张典、赵晓晴、李建廷、叶先敏、孙子翔、王小宝。

ARXIV

基于智能体的流动性与能源政策建模：以2050年芝加哥地区为案例
大都市区正同步推进多项干预措施，以提升交通流动性、可达性及能源效率，亟需集成化工具来理解这些政策如何相互作用，进而影响出行行为、能源消耗及基础设施需求。本文以‘照常情景’（BAU）为基准，评估电气化、货运需求管理、道路收费、停车改革及公共交通扩展等政策组合对2050年芝加哥大都市交通系统的综合影响。我们采用POLARIS——一个基于2019年条件校准的大规模智能体建模框架，在伊利诺伊州东北部七县区域模拟九种政策情景。

近期研究重心从静态街景图像的单帧感知建模，转向关注感知的动态性、主体性与行为闭环——尤其强调局部视觉输入如何支撑真实尺度城市中的空间定位、导航与鲁棒决策。方法上正融合MLLM、三维地理仿真与人类感知机制（如注视）以弥合感知与行动鸿沟。

近30天 15 近7天 5 来源 41 论文 203

趋势信号

- UrbanGround提出首个基于全境三维地理空间数据构建的物理约束式城市仿真沙盘，支持第一人称视角闭环交互与导航任务评估
- PairWise Image Finder工具显式引入语义分割掩膜量化街景图像对的视觉对齐质量，凸显纵向变化分析中视角与对齐精度的关键性
- Place Pulse-Gaze数据集首次同步整合街景图像、眼动追踪记录与个体感知标签，推动人类注视建模成为主观城市感知的新基准信号
- 多篇论文集中检验街景感知结果在天气条件、平台差异、路径可用性变化及行人干扰下的稳定性与偏差

核心观点

- 街景图像的感知建模正经历从‘客观属性估计’（如生理唤醒）向‘主观判断形成机制’（如注视引导）的范式延伸
- 局部街景感知本身存在固有局限：其价值高度依赖后续空间行为能力（如定位→导航→鲁棒响应），不能脱离移动智能体的具身闭环进行评估
- 跨平台、跨时间、跨天气的街景感知结果一致性尚未建立，测量偏差已成为影响城市感知可比性与政策适用性的核心方法论问题
- 多模态大语言模型（MLLM）被寄望于整合视觉、地理与语言信息以实现空间主体性，但其在真实尺度三维城市中的行为边界仍需严格实证检验

RESEARCH IDEA

MLLM街景感知在真实尺度导航中失效于路径可用性突变

UrbanGround中基于街景输入的MLLM智能体在路径可用性发生突变（如临时封路、施工围挡）时无法实时修正导航策略，因其视觉-动作映射未建模物理约束的即时可观测性缺口

为什么现在值得做：城市更新高频化与应急管理需求使实时路径可用性成为智慧城市操作系统的基础信号；UrbanGround仿真框架已开源并支持闭环交互，为注入动态约束提供了可控实验接口。

关键难点

- 需定义‘路径可用性突变’在街景图像中的可观测代理特征（如围挡纹理密度、施工标识语义掩膜覆盖率
- 需将物理约束（如通行宽度阈值、坡度限制）编码为MLLM可解析的结构化提示，而非仅依赖视觉token
- UrbanGround当前不提供动态障碍物API接口，须扩展其仿真引擎以支持按需触发状态变更

建议切入

- 在UrbanGround中构建三类典型突变事件（临时封路、占道施工、人行道中断），通过PairWise Image Finder验证其街景图像视觉对齐质量，确保变化可被检测
- 基于Place Pulse-Gaze数据集提取注视热点与路径关键区的空间重叠率，量化MLLM注意力是否自然聚焦于可用性判别区域
- 设计约束感知提示模板：将OSM道路属性（surface、width、barrier）与街景分割掩膜（如‘construction-barrier’类）联合嵌入MLLM上下文，测试其重规划响应延迟

REPRESENTATIVE ITEMS

TRANSACTIONS IN GIS

利用街景图像估计客观城市感知：以生理唤醒为例

《国际地理信息系统汇刊》（Transactions in GIS），第30卷，第5期，2026年8月。

COMPUTERS, ENVIRONMENT AND URBAN SYSTEMS

深度学习揭示街景影像在城市感知制图中的跨平台一致性

出版日期：2026年12月；来源：《Computers, Environment and Urban Systems》，第130卷；作者：白婷婷、徐东、高峰、刘佳伟、宋永泽。

COMPUTERS, ENVIRONMENT AND URBAN SYSTEMS

天气有影响吗？基于街景图像的城市感知评估中的测量偏差探究

发表日期：2026年7月 来源：《计算机、环境与城市系统》，第127卷 作者：金东焕，李承敏，韩彩妍，金友贞，高奉宇，黄义正。

ARXIV

UrbanGround：从局部感知到真实尺度城市中的空间主体性

多模态大语言模型（MLLM）能够解析街景图像，但城市主体性取决于此类局部感知证据在智能体开始移动后是否仍具实用性。本文探究当前MLLM智能体在复杂的真实尺度城市中，将局部城市感知转化为可靠行为的能力边界。我们提出UrbanGround——首个基于全境三维地理空间数据构建的香港物理约束式仿真沙盘，使该问题得以实证检验。

UrbanComp Lab 学习资料库 (https://research.urbancomp.dev/)

USDOT OPEN DATA

美国国家公路交通安全管理局缺陷调查办公室 (NHTSA-ODI)

录入美国国家公路交通安全管理局缺陷调查办公室 (NHTSA-ODI) 车辆车主投诉数据库的投诉信息，将与其他数据源结合使用，以识别需开展调查的安全问题，并判断是否存在与安全相关的缺陷趋势。此外，投诉信息还用于监测现有召回措施的覆盖范围是否适当、补救是否充分。

USDOT OPEN DATA

车辆安全研发数据库——基于车辆屏障参数的车辆测试查询

研发数据库 (R&D Database) 提供车辆碰撞测试数据、生物力学测试数据及零部件测试数据，以支持美国国家公路交通安全管理局 (NHTSA) 的机动车与交通安全目标。

TRANSPORTATION RESEARCH PART C

F2V-Former

出版日期：2027年1月；来源：《交通运输研究C辑：新兴技术》，第194卷A期；作者：周乐琪、金建刚、张艳飞、闫然。

CITIES TODAY

新实验室联合城市测试人工智能工具

地方政府技术提供商CivicPlus已成立新部门，致力于开发和 本文首发于 Cities Today。

USDOT OPEN DATA

CAFE (企业平均燃油经济性) ——灵活燃料车辆/AMFA

美国国家公路交通安全管理局 (NHTSA) 的企业平均燃油经济性 (CAFE) 计划要求在美国销售的乘用车及轻型卡车制造商达到以英里每加仑 (mpg) 表示的CAFE标准。该计划旨在通过提升乘用车及轻型卡车的燃油经济性，降低全国能源消耗。CAFE公共信息中心 (PIC) 是CAFE计划数据的权威来源，用户可通过该网站以报告和/或图表形式查看燃油经济性数据；数据支持排序与筛选，以生成定制化报告，并可下载为Excel或PDF格式文件。NHTSA定期更新PIC中的CAFE数据，因此每份报告与图表均标注日期，以表明NHTSA最近一次更新时间。

ARXIV

EDGE：基于图结构领域特定语言配置的对话模拟确定性图评估引擎

随着智能体系统演进为复杂的多智能体协同工作流，对智能体行为一致性与确定性的系统化评估框架的需求日益迫切且关键。本文提出一种形式化评估方法，该方法以 AgentGraph 为基础——这是一种由领域特定语言 (DSL) 驱动的规划器，通过可动态调整的有向图表征智能体推理过程。我们利用该结构化形式化表示，结合图遍历算法穷举枚举所有可能的对话路径，从而构建一个全面的评估集，覆盖智能体的全部行为空间。

ARCHDAILY

亚鲁莫公寓 / 特罗皮科建筑事务所 (Trópico Arquitectura)

一套为年轻夫妇设计的公寓，旨在优化财务资源，同时充分利用建筑结构的空间特性。

ARXIV

面向更快工具调用智能体的推测性宏提交 (Speculative Macro

使用工具的大型语言模型 (LLM) 智能体不仅在模型推理上消耗真实时间，还需经历串行的动作—观测轮次：每次工具调用、环境状态转移及观测结果均会延迟后续决策。我们提出 \textbf{推测性宏提交} (Speculative Macro Commit, SMC)，一种面向双层智能体系统的运行时机制：一个大型权威执行者模型 (authoritative actor model) 生成正式轨迹，而一个更快的推测性起草者模型 (speculative drafter model) 则持续在一个隔离的环境快照上预测并执行未来动作链。