

# 基于多模态融合的智慧街镇基层治理平台

陈文, 梁华宇, 廖红虹\*

联通(广东)产业互联网有限公司, 广东 广州

收稿日期: 2025年5月6日; 录用日期: 2025年6月6日; 发布日期: 2025年6月13日

## 摘要

数字技术和人工智能技术赋能基层治理是适应基层政府数字化转型发展的必然要求, 也是提升基层治理能力现代化水平的重要途径。本文提出了一种基于多模态数据融合的智慧街镇基层治理平台, 旨在通过整合多源数据, 构建高效、智能的基层治理新模式。该平台以多模态融合的基层治理大模型为底座, 结合大小模型协同技术, 实现了事件的自动填报、精准派发和智能处置, 有效提升了基层治理的自动化和智能化水平。通过多源设备联动和应用迭代, 平台推动了基层治理从模糊到精准、从被动应对到主动研判的转变, 为街镇基层治理提供了有力的技术支撑。

## 关键词

基层治理, 多模态大模型, 大小模型协同, 多源数据融合

# Smart Street and Town Management Platform Based on Multimodal Fusion

Wen Chen, Huayu Liang, Honghong Liao\*

China Unicom (Guangdong) Industrial Internet Co., Ltd., Guangzhou Guangdong

Received: May 6<sup>th</sup>, 2025; accepted: Jun. 6<sup>th</sup>, 2025; published: Jun. 13<sup>th</sup>, 2025

## Abstract

Empowering grassroots governance with digital technology and artificial intelligence technology is an inevitable requirement for adapting to the digital transformation and development of grassroots governments, and is also an important way to improve the modernization level of grassroots governance capabilities. This paper proposes a smart street and town management platform based on multimodal data fusion, aiming to build an efficient and intelligent new model of grassroots governance by integrating multi-source data. The platform is based on a large model of grassroots governance with

\*通讯作者。

文章引用: 陈文, 梁华宇, 廖红虹. 基于多模态融合的智慧街镇基层治理平台[J]. 计算机科学与应用, 2025, 15(6): 24-34. DOI: 10.12677/csa.2025.156154

**multimodal fusion, and combines large and small model collaboration technology to achieve automatic reporting, accurate distribution and intelligent handling of events, effectively improving the automation and intelligence level of grassroots governance. Through the linkage of multi-source equipment and application iteration, the platform has promoted the transformation of grassroots governance from fuzzy to precise, from passive response to active research and judgment, and provided strong technical support for grassroots governance in streets and towns.**

## Keywords

**Grassroots Governance, Large Multimodal Models, Large and Small Model Collaboration, Multi-Source Data Fusion**

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着城市化进程的加速，街镇基层治理作为国家治理体系的末端，面临着复杂多样的管理需求。传统的治理模式已难以满足现代社会对高效、精准治理的要求。街镇基层治理能力因其治理对象、治理资源、治理结构和治理机制的不同而与国家治理能力存在显著差异。街镇基层治理的主要任务是处理与民众生活息息相关的“小事”，而非完成“大事”，但越到基层，事情越复杂，制约条件越多。街镇基层治理具有以下特点[1]：

- 1) 任务繁杂：基层治理工作涉及党群工作、社区建设、综治安全、城市管理、民生保障等多个方面，呈现出“上面千条线，下面一根针”的局面。
- 2) 工作量大：基层治理任务繁重，责任重大，基层干部常常需要“5+2”“白+黑”地工作，甚至“星期六保证不休息，星期天不保证休息”成为工作常态。
- 3) 资源匮乏：街镇基层组织拥有的资源有限，无论是资金还是技术都严重缺乏，导致治理工作“既难治又难理”。
- 4) 成就感低：基层事件难以预防，治理攻坚行动限时完成，基层管理者忙于应对，缺乏成就感。

2021年7月，中共中央、国务院印发《关于加强基层治理体系和治理能力现代化建设的意见》，明确指出要“加强基层智慧治理能力建设”，并从“规划建设”“整合数据资源”“拓展应用场景”等方面对基层社会治理数字化转型提出了改革方向。在此背景下，数字技术与人工智能(AI)的结合为智慧街镇管理提供了新的思路和方法[2][3]。同时，新时期的“枫桥经验”作为基层治理的典范[4][5]，其创新发展为智慧街镇管理平台的建设提供了宝贵经验。具体体现如下：

- 1) 群众参与：智慧街镇管理平台通过多模态数据融合，整合居民反馈、社区活动等信息，鼓励居民参与治理，形成人人有责、人人尽责的治理共同体。
- 2) 矛盾化解：利用AI技术对居民诉求进行语义分析和分类处理，实现矛盾纠纷的快速响应和精准化解。
- 3) 资源整合：通过大数据和物联网技术，实现街镇内人、地、物、事、情、组织等基础信息的整合，提升治理的精细化水平。

然而，区县和街镇级别政府的各行业数据通常分布在不同的部门和机构中，数据孤岛现象严重，导致数据管理、数据融合及业务协同能力不足。同时，在城市治理和事件处置中，传统的事件感知方式滞

后,难以及时准确地响应城市管理问题[6]。因此,迫切需要构建一个基于数字技术和人工智能技术的集事件感知和处置为一体的综合基层治理系统,实现街镇基层治理事件的实时感知、精准分析、整体研判及协同指挥。

## 国内外研究现状

近年来,数字技术与人工智能(AI)在街镇基层治理中的应用呈现爆发式增长,成为破解传统治理难题、提升精细化水平的关键抓手。日本 2016 年提出了社会 5.0 战略,强调通过智能化技术解决偏远街镇的交通、医疗等难题。欧盟 2017 年启动了“欧盟智慧乡村行动”,搭建了包含“基础设施”“技术平台”“数字服务”与“社会应用”的智慧街镇乡村的数字生态系统,为农村居民提供了包括数字村庄平台、“爱心村”、信仰平台、数字教育等一系列智能服务,有效提升了街镇基层治理现代化水平[6]。

几年来,国内各地普遍构建“数据中枢 + 场景应用”的治理架构,此类平台的核心逻辑在于数据融合驱动 AI 辅助决策优化。如上海静安寺街道的“水管家”智能水表通过动态监测独居老人用水异常,将被动救助转为主动预警,覆盖 2.3 万名高危群体[7]。杭州西湖区的“镇街大脑”整合数字党建、平安、治理等六大模块,实现事件“一码流转、闭环处置”,并通过“吹哨报到”机制提升跨部门协同效率[8]。北京市海淀区在党建引领基层治理中推出的创新试点项目“社易智”智能助手,旨在通过人工智能技术优化社区服务流程,提升治理效能。该助手充分利用人工智能大模型技术及 APP、小程序等线上方式,整合服务资源,畅通社会参与渠道,通过数智化手段推动基层治理减负增效。2024 年 1~8 月,海淀区共承办群众诉求 119.4 万件,诉求响应率保持 100%,解决率由 2023 年的 93.49% 上升至 97.74%,满意率由 94.97% 提升至 97.99% [9]。

本文提出的“基于多模态融合的智慧街镇基层治理平台”是针对街镇基层治理业务需求研发的一套基于多模态融合的智能化平台,旨在通过人工智能技术为街镇治理提供全面业务支撑。该平台研发了: 1) 多源数据整合与共享的高效方法[10],构建了镇域数据中枢、业务中枢和智能中枢,实现全科网格治理和智能化事件处理流程。2) 集成了以基于华为 MindSpore 框架[11][12]的 AI 视觉识别算法和基层治理大模型为核心的智能中枢,提供数十种基于复杂环境下的任务识别算法和基层治理垂域大模型; 3) 面向街镇用户提供创新应用,结合大小模型协同技术[13],有效提升了基层街镇管理问题的自动识别和处理能力。平台通过多源设备联动和基于多模态数据融合的各类应用,实现基层治理从模糊到精准、从被动应对到主动研判转变。以大小模型协同的多模态 AI 能力为核心,实现在复杂环境下高精度事件自动填报和处置,有效提高街镇基层治理的自动化处理能力。

最近范建等[14]研究了人工智能(AI)和大数据技术在智慧城市基层治理中的应用,通过结合大数据、NLP、深度学习神经网络等前沿技术,设计并实现了一套高效的基层治理平台。该平台利用卷积循环神经网络(CRNN)提高图像识别精度,通过大数据 ETL 技术构建多样化数据集,以及实现网格管理、事件管理和全景分析等多个核心业务模块的智能化升级。本文提出的基于多模态融合的智慧街镇管理平台主要创新点在于多模态数据融合与大小模型的协同,实现从多源数据汇聚到多模态识别的一体化能力,在实现过程中充分吸收了国内外先进的数据治理方法与人工智能尤其是大模型技术能力,在技术上通过构建多模态基层治理垂域大模型[15]-[18]及其创新应用,实现对基层治理的效率增加与使用的便利性。

本文内容结构如下:第 2 节主要介绍平台的技术架构与技术创新,重点介绍智慧街镇基层治理平台的技术架构与特点、基于联通元景多模态大模型的基层治理垂域大模型及其模型微调与应用部署,基于大小模型协同技术的面向街镇用户创新应用:AI 自动填报精灵和事件预案 + 派发引擎;第 3 节主要介绍相关成果,包括技术成果和商业成果。最后,对全文进行总结。

## 2. 智慧街镇基层治理平台介绍

### 2.1. 平台技术架构

智慧街镇基层治理平台围绕基层政府的工作规划，整合街镇治理的多数据源，构建起“镇域一网统管”与“全科网格治理”紧密结合的智能化治理体系，落实事件快速处置机制，助力构建“全域统筹、多源联动、网格治理”的基层治理格局。平台以打造街镇业务闭环、实现数据本地留存、横向打通信息化孤岛和壁垒为目标，在多源数据汇聚、AI 智能算法和面向街镇用户创新应用多个层次技术创新，其核心架构图如图 1 所示，总体而言，架构具有以下特点：

(1) 多源数据汇聚：当前，街镇基层治理面临着多系统数据分散的问题，各委办局自建应用和行业数据之间缺乏有效的整合与协同，导致数据资源无法充分发挥其价值。为解决这一问题，本文系统聚焦于数据的拉通与汇聚，通过整合各委办局自建应用和行业数据，并结合省市回流数据，构建街镇基层治理的完整数据底数。在此基础上，建设镇域数据中枢，实现“多网融合、一网统筹”，为基层治理提供全面、精准的数据支持。数据汇聚的起点是基层工作人员所使用的各类硬件设备，这些设备是数据采集的重要源头。本文系统从视频监控、无人机、门禁系统、水电表等设备入手，通过标准化接口和协议，实现数据的互融互通。这些设备采集的数据涵盖了社会治理的多个方面，包括公共安全、城市管理、民生服务等，是基层治理的重要依据。通过打破数据壁垒，我们将这些分散的数据源汇聚到镇域数据中枢，形成全域数据的多源汇聚。在数据汇聚过程中，本文系统注重数据的质量和安全性。通过数据清洗、校验和加密等技术手段，确保汇聚的数据准确、可靠且符合安全标准。同时，建立数据更新机制，确保数据的时效性，使其能够真实反映基层治理的动态变化。通过多源数据汇聚，不仅为基层治理提供了丰富的数据资源，还为后续的数据分析、决策支持和智能化应用奠定了坚实的基础。

(2) 共性能力集约：通过集约化建设，能够整合各类基础能力，形成高效协同的治理模式。在 AI 能力方面，构建基于 MindSpore 框架国产化的视觉识别算法与基于元景多模态大模型的基层治理大模型能力服务。MindSpore 框架作为国内领先的 AI 开发平台，提供了强大的计算能力和灵活的开发环境，能够为基层治理提供智能化的算法支持。而元景多模态大模型则具备处理多种数据类型的能力，包括文本、图像和视频等，能够全面感知和理解基层治理中的各类场景和事件。在此基础上，通过大小模型的协同赋能，进一步提升对图像和视频数据的识别与分析能力，赋能基层各应用系统能够实现智能化升级，从而更好地服务于基层治理的各个环节。最终，所有这些能力将汇聚于一张大屏之上，实现集中监控与一网呈现。这张大屏不仅是信息的展示窗口，更是基层治理的指挥中枢，通过它，基层治理人员可以实时掌握各类事件的动态，快速做出决策。

(3) 创新应用承载：建设一系列创新应用，覆盖政务、民生等多个方面，打造本地横向业务闭环和数据流通。同时，优化本地网格基础信息维护、事件上报及处置体系，创新推出了基于基层治理大模型的 AI 自动填报精灵和事件预案及派发引擎，打造街镇级事件处置智能中枢。AI 自动填报精灵可识别网格员上传的图片或视频，自动匹配数百类网格事件并生成事件描述，减少手动操作，提升工作效率与质量。事件预案及派发引擎则在事件处置前自动从预案库中匹配应对策略，并向责任部门派发工单，确保紧急情况下快速高效应对。事件预案及派发引擎通过持续分析历史事件数据和处置结果，不断调整和优化预案库中的策略，使其能够更精准地应对各种复杂多变的网格事件。同时，该引擎支持多部门协同工作，能够根据事件的性质和紧急程度，自动协调涉及的多个责任部门，打破部门之间的信息壁垒，实现信息共享与协同处置。此外，它还具备实时监控功能，能够对事件处置过程进行全程跟踪，及时发现并纠正偏差，确保事件能够按照既定预案高效推进。通过这些功能的综合运用，事件预案及派发引擎不仅能够提升事件处置的效率和准确性，还能为网格化管理提供更加科学、智能的决策支持，助力街镇基层治理

迈向精细化、智能化的新阶段。



Figure 1. Technical architecture for the platform  
图 1. 平台技术架构图

## 2.2. 平台创新技术介绍

### 2.2.1. 基层治理大模型



Figure 2. Technical architecture of multimodal domain large model for grassroots governance  
图 2. 多模态基层治理垂域大模型架构

本文提出的基层治理大模型是基于中国联通元景多模态大模型底座，创新性提出了 MOE LoRA 模型架构[19]-[22]，结合多模态(文本、语音、图片)的街镇基层治理数据对底座大模型进行微调，构建满足街

镇基层治理需要的多模态垂域大模型。该基层治理大模型通过综合处理和分析这些多模态数据，为基层治理提供全面、高效、智能化的决策支持。模型的架构如图 2 所示。

本文提出的多模态基层治理垂域大模型具备的基本能力：

(1) 多模态数据融合与理解。该模型能够处理多种数据类型，包括文本、图像、视频、音频等，通过街镇对接街镇视频监控中心，可以分析监控视频中的图像和音频数据，识别异常行为或事件；可以对文本数据进行深度语义理解，能够处理自然语言描述的事件报告、居民诉求等信息，提取关键信息并进行分类；结合多源数据汇聚，可以将不同模态的数据进行关联分析，例如将居民的投诉文本与对应的地理位置、监控视频或传感器数据相结合，形成更全面的事件描述。

(2) 事件识别与分类。通过分析多模态数据，自动识别各类基层治理事件，如社区安全事件、环境问题、交通拥堵等；根据事件的性质、严重程度和紧急性，将事件分类并标记优先级，为后续的处置提供依据；实时监测事件的发展动态，通过视频监控或传感器数据，及时发现事件的演变趋势。

(3) 智能决策支持。根据事件类型和历史数据，自动匹配预设的处置预案，为责任部门提供标准化的处置建议；结合事件的地理位置和资源分布，智能调度基层工作人员、设备和物资，确保快速响应；通过数据分析，评估事件可能带来的风险和影响，为决策提供科学依据。

同时，针对本文提出的多模态基层治理垂域大模型在训练方面所面临的泛化能力不足和过拟合难题。在训练过程中，本文采用了模型迁移学习结合数据增强的训练方法，使用在大规模多模态数据集上预训练的联通元景 34 B 大模型，通过同义词替换、回译、随机删除、随机交换等多种文本数据增强方法的组合和图片的各种增强操作方式，生成更加多样的训练样本，解决小数据量大模型训练难题。迁移学习过程则是基于预训练的基础大模型，使用标注好的多模态基层治理数据进行 LoRA 方式的微调，让新的模型学习到已有模型的丰富特性，解决泛化能力不足的问题，组合多种数据增强方法让训练数据更丰富，为模型提供基层治理场景中的高质量训练数据，确保模型具备足够的基层治理应对能力。

多模态基层治理垂域大模型的微调训练和应用服务架构如图 3 所示，为确保用户数据的本地保存及数据安全，在用户域和模型域之间，设置了转发域，即只有在开通白名单的应用才可以对接模型，由网管和负载均衡任务确保模型的稳定性和数据的安全性。

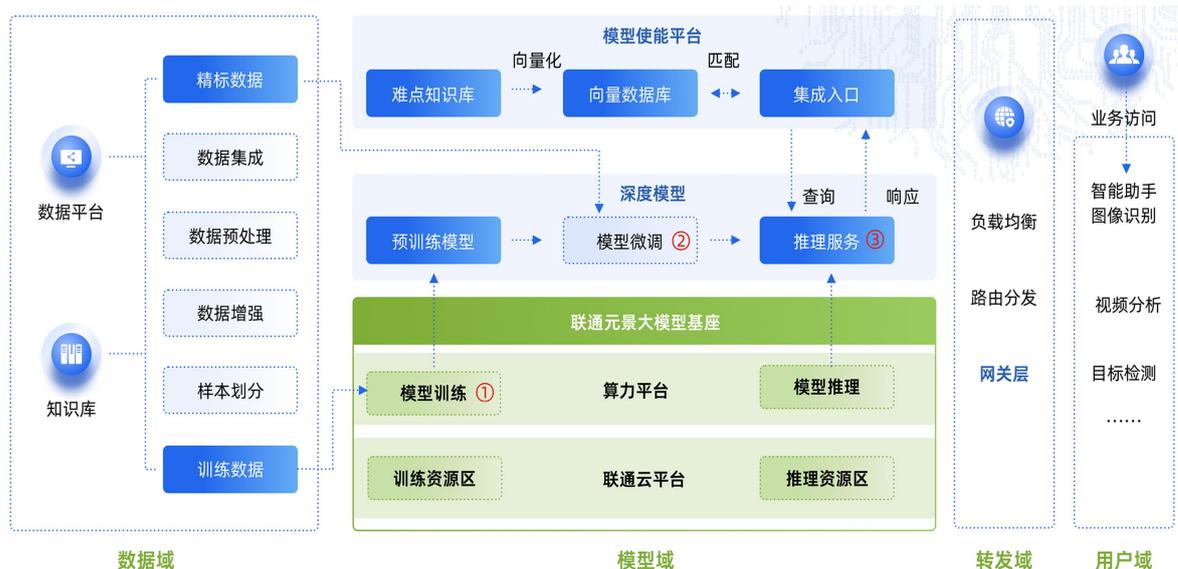


Figure 3. Model optimization and application services

图 3. 模型调优与应用服务

在实际应用过程中，本文采用基于 RAG 框架[23]的知识管理和业务应用系统架构，以业务应用、知识库和多模态基层治理垂域大模型三部分组成，具体如图 4 所示。知识库管理负责上传和维护各种格式的文档，通过嵌入模型[24]-[26]处理多模态数据后存储在预料知识库、事件特征知识库等多层次知识库中。业务应用部分，用户通过前端和后端系统提交请求，经由嵌入模型向量化处理后，使用 Milvus 向量数据库[27]进行检索，筛选和优化答案后返回给用户。联通元景大模型底座通过数据收集、标注和存储、知识更新与冲突自动发现，结合基础模型和增量模型进行数据处理和分析。整个架构的创新点在于 RAG 框架的应用、嵌入模型的使用、Milvus 向量数据库的高效检索、多层次知识库的结构化管理以及全流程自动化的数据处理能力。

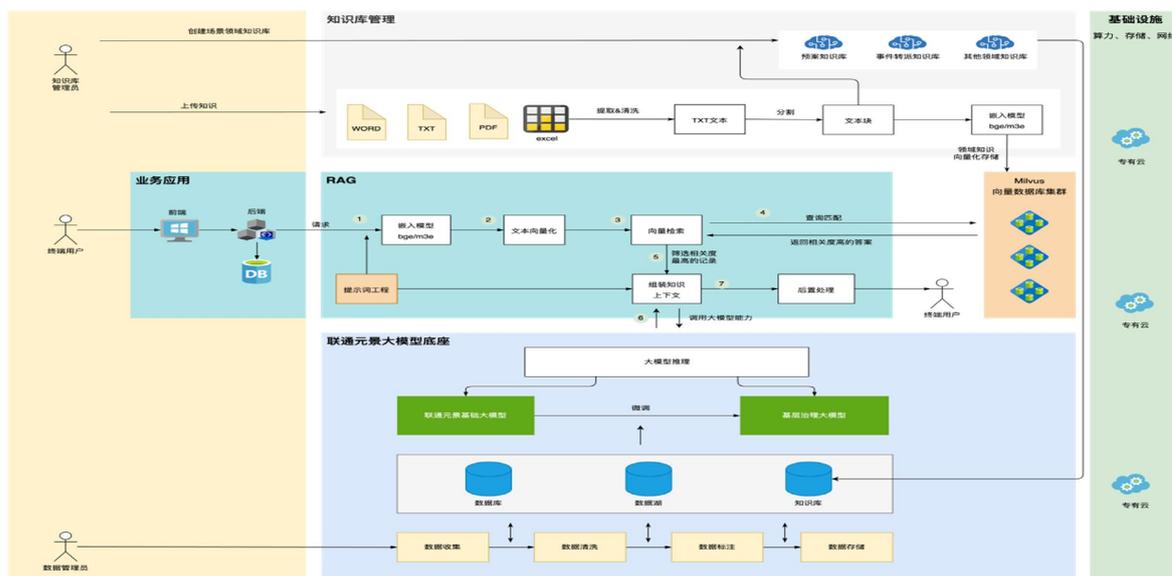


Figure 4. Technical architecture of LLM application for grassroots governance  
图 4. 基层治理大模型应用技术架构

### 2.2.2. 基于大小模型协同的街镇创新应用

基于前文的多模态基层治理垂域大模型，结合以华为 MindSpore 框架的 AI 视觉识别算法，构建基于大小模型协同的街镇创新应用，实现基层事件“发现、派发、处置、跟踪、归档”的全流程管理，快速精确打通基层事件触达和处置的最后一公里。基于此，本文创新推出了基于基层治理大模型的 AI 自动填报精灵和事件预案及派发引擎，打造街镇级事件处置智能中枢。

#### (1) AI 自动填报精灵

AI 自动填报精灵主要解决基层网格员日常巡检工作中的繁重工作，在传统的工作模式中，网格员上报一个事件需要在数百个事件类别中选择一个，并上传事件现场照片和手工填写事件内容，在实际操作过程中，至少需要 10 分钟以上才能完成一个事件的申报。这就导致很多事情网格员事件申报不及时，甚至不愿意申报的情况。

AI 自动填报精灵采用大小模型协同技术，在网格员巡视过程中，发现事件存在时，通过拍摄现场视频或者照片上传，由多模态基层治理垂域大模型对上传视频或图片做初步识别到大体事件的分类，再调用 AI 视觉识别算法做进一步精确事件识别，并由大模型根据识别结果生成事件内容描述，最终人工仅需进行核验即可完成事件填报。据统计，整个申报过程仅需 30 秒，人工操作 2 次即可完成事件申报。当发现 AI 模型识别出现偏差时，还可以通过语音输入的方式对事件描述内容和事件类别进行二次矫正。

AI 自动填报精灵的详细实现过程如图 5 所示。

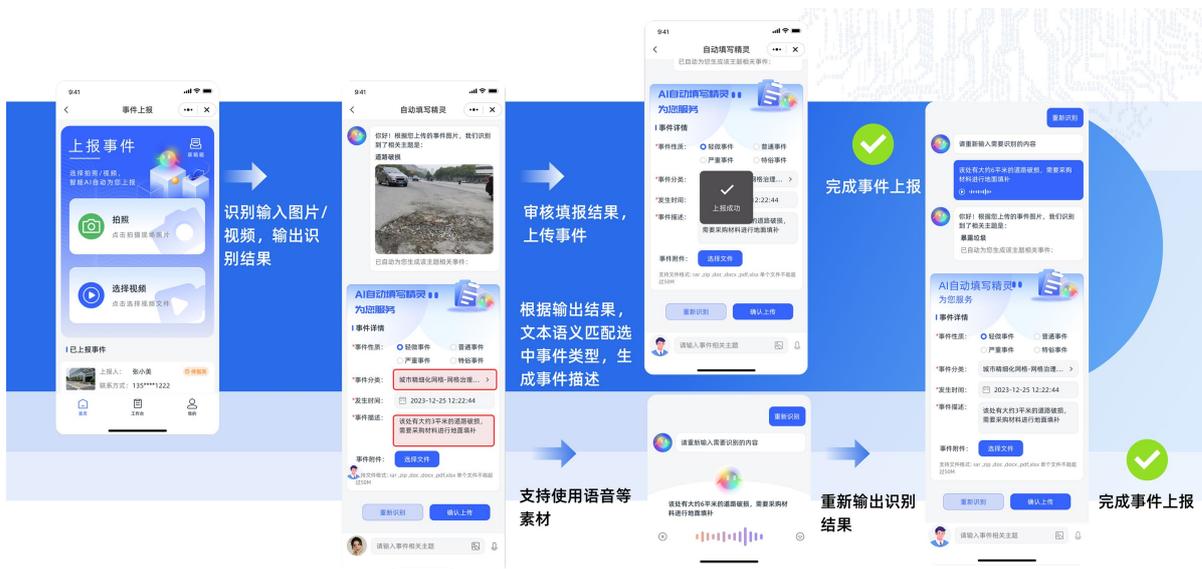


Figure 5. Flow of AI automatic reporting wizard  
图 5. AI 自动填报精灵实现流程

(2) 事件预案及派发引擎

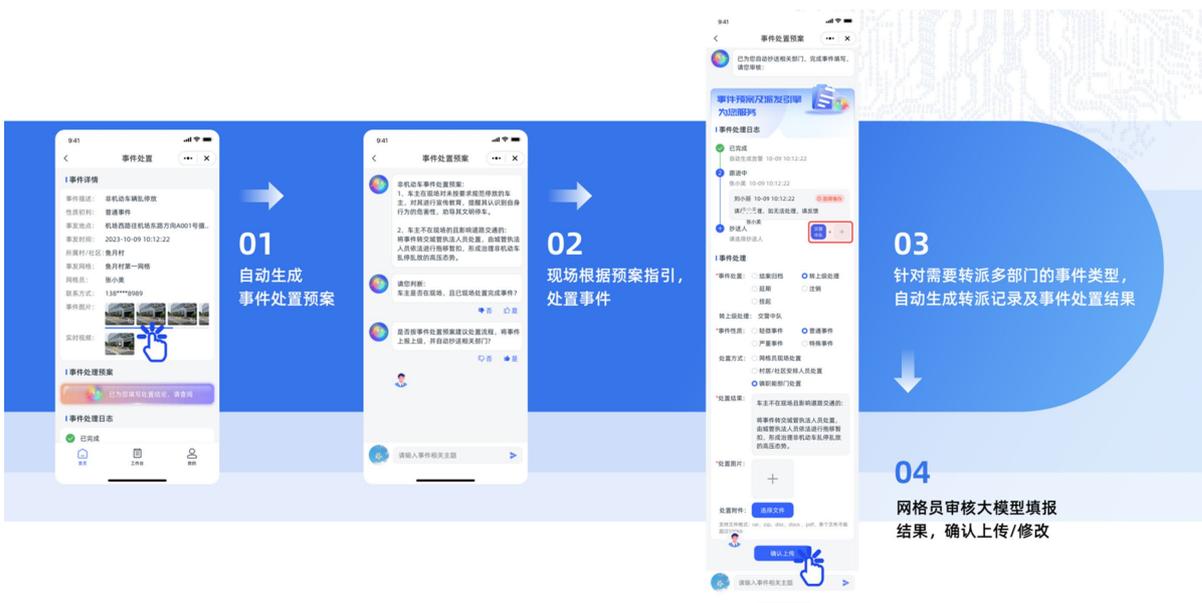


Figure 6. Flow of event planning and dispatching  
图 6. 事件预案与派发实现流程

事件预案及派发引擎主要是解决基层管理人员流动性大，急需各类处置预案指导工作，同时，基层事件常常涉及多个部门，事件流转难以找到对应的接口人。传统的处置模式是，在接到事件告警工单后，先通过电话等方式寻求事件处置指导，这通常会涉及到多个部门，需要跟多个部门进行多番沟通后才能明确事件处置的下一步流程。整个沟通过程费时费力，且对基层办事人员来说毫无成就感可言。

事件预案及派发引擎通过对事先梳理好的事件处置预案，并将事件处置预案存储在 RAG 框架的知识库中，在实际应用过程中，对于填报上来的事件工单，采用多模态基层治理垂域大模型对事件工单进行预案精准匹配，并形成快速指导方案，同时，对事件工单进行自动派单建议，在人工确认后即可完成整个工单派发任务，落实工单处置责任。同时，可以根据事件工单派发情况，对事件预案进行更新，保持事件预案的动态更新，实现事件预案“建设一次，动态更新”。

事件预案及派发引擎的详细执行流程如图 6 所示。

### 3. 技术与商业成果

本节主要介绍本文提出的智慧街镇管理平台在技术上和商业上取得的成果。

#### 3.1. 技术成果

##### 3.1.1. 数据整合与业务协同提升

平台的镇域数据中枢通过整合来自不同政府部门和行业的数据，实现了各委办局自建应用的互通互联，形成了统一的数据管理平台。以基于本份解决方案实施的台山市新型智慧城市项目为例，项目数据中枢已接入 25 个市直单位 22 个业务系统平台，约共 1428 类 1.83 亿条数据。集成的数据通过数据中枢治理后，形成了 690 类开放库数据资产。平台成功拉通了市、区、镇各层级的多源数据，实现了“多网融合、一网统筹”的数据协同。这不仅显著提升了数据管理的效率，还促进了各部门的业务协同，减少了以往因信息孤岛问题带来的沟通延迟。

##### 3.1.2. 事件响应速度和精准度显著提升

在网格事件的处理过程中，平台通过 AI 自动填报精灵和事件预案及派发引擎，大幅减少了网格员手动输入的工作量。基于大模型的自动化事件描述和预案生成功能，基层事件处理速度提高了约 2000%。同时，事件处置的精准度得到了显著提升，AI 识别技术在复杂场景下的识别准确率达到了 95% 以上。这一技术突破极大地提升了基层事件响应的时效性。

##### 3.1.3. 智能巡检与多源联动的成功应用

平台整合了固定检测点、无人机巡检等多源数据，通过 AI 智能识别技术，成功实现了对城市管理事件的实时监测和自动上报。在珠海市项目中，系统每日处理的事件数量较传统模式增加了 40% 以上，且通过自动化分析和分派功能，事件的处理时间缩短了近 50%。智能巡检的应用不仅减少了人工巡查的工作量，还大幅提升了事件的处理效率和覆盖率。

#### 3.2. 商业成果

在市场推广上，通过提供标准化的解决方案和开放接口，本文系统实现了强大的可复制性和可扩展性。方案的成功落地带来了显著的商业回报，在台山市和珠海市项目合同金额均达到 2000 万元以上，累计服务客户超 20 家，本文系统获得了超亿级智慧街镇项目收入。此外，方案还支持开放模型应用接口的模型进行功能调用，极大降低了基层政府的信息化建设成本，进一步提升了项目的可持续发展能力。

#### 3.3. 未来的优化与扩展方向

平台在现有成功应用的基础上，未来将继续优化 AI 大模型的性能，进一步提升系统的识别精度和事件处理能力。同时，将继续探索平台在更多基层治理场景中的应用，例如智慧社区、智慧园区等，扩展平台的应用范围，充分发挥其在数字化治理中的潜力。通过不断积累基层治理数据，平台的大模型也将得到进一步优化，帮助基层政府应对更加多样化的治理需求。

## 4. 总结

本文提出的基于多模态融合的智慧街镇管理平台，通过整合多源数据、构建基层治理大模型以及创新应用开发，为街镇基层治理提供了一种高效、智能化的解决方案。平台在技术架构上实现了多源数据汇聚、共性能力集约和创新应用承载，有效解决了传统基层治理中存在的孤岛、资源匮乏、效率低下等问题。通过多模态数据融合与大小模型协同，平台在事件自动填报、事件派发机制和预案自主处置等方面表现出显著优势，极大地提升了基层治理的自动化处理能力。在商业成果方面，平台凭借其标准化的解决方案和开放接口，展现出强大的可复制性和市场竞争力，获得了显著的商业回报，并为基层政府的信息化建设提供了低成本、高效能的解决方案。

然而，随着应用场景的复杂化和多样化，平台在特定极端场景中的识别和处理能力仍需进一步优化。未来的研究将聚焦于提升平台对不同场景的适应能力，通过不断积累基层治理数据和优化模型，进一步推动智能基层治理的广泛应用。同时，平台将继续探索在更多基层治理场景中的应用，如智慧社区、智慧园区等，扩展其应用范围，充分发挥其在数字化治理中的潜力。

本文的研究成果表明，利用 AI 大模型能够显著提升基层治理的效率和精准度，特别是在事件处理上，本文表现出了显著的优势。通过成功的商业推广，平台展示了其高效的复制性和市场竞争力。未来的研究将致力于进一步优化平台性能，拓展其应用领域，为基层治理的数字化转型提供更有力的支持。

## 致 谢

感谢各地方政府和基层治理部门的工作人员，在平台的试点和应用过程中，积极配合数据收集与反馈，为平台的实际验证提供了重要的帮助。

## 参考文献

- [1] 王海侠. 基层社会治理存在的十大突出问题[EB/OL]. 经济导刊. [http://www.qstheory.cn/qshyix/2024-09/18/c\\_1130202758.htm](https://www.jingjidaokan.com/icms/null/null/ns:LHQ6LGY6LGM6MmM5ZTg0NzU4Y2EwZTU4YjAx-OGU5ZGExY2MwZjAwNDlscDosYTosbTo=/show.vsm], 2025-03-05.</a></li>
<li>[2] 求是网. 以数字技术赋能基层治理[EB/OL]. <a href=), 2025-03-05.
- [3] 司海燕. 推动人工智能赋能基层治理[EB/OL]. 新华日报. [https://theory.jschina.com.cn/sxzk/xl/zxtj/202502/t20250226\\_8460582.shtml](https://theory.jschina.com.cn/sxzk/xl/zxtj/202502/t20250226_8460582.shtml), 2025-03-08.
- [4] 吴锦良. “枫桥经验”演进与基层治理创新[J]. 浙江社会科学, 2010(7): 43-49, 29, 126.
- [5] 中华人民共和国中央人民政府. 谱写基层善治新篇章——坚持和发展新时代“枫桥经验”综述[EB/OL]. [https://www.gov.cn/yaowen/liebiao/202311/content\\_6913690.htm](https://www.gov.cn/yaowen/liebiao/202311/content_6913690.htm), 2025-03-08.
- [6] 冯献, 李瑾, 崔凯. 乡村治理数字化: 现状, 需求与对策研究[J]. 电子政务, 2020(6): 73-85.
- [7] 李友梅, 黄晓春. 上海城市数字化治理现状及推进思路[J]. 科学发展, 2022(12): 5-13.
- [8] 银江技术股份有限公司. 镇街大脑[EB/OL]. [http://www.enjoyor.net/solution\\_judicial.html](http://www.enjoyor.net/solution_judicial.html), 2025-03-05.
- [9] 海淀统战. 海淀: 深化“大家商量着办”践行“围着基层转”共建美好家园[EB/OL]. [https://paper.people.com.cn/rmr/pc/content/202505/22/content\\_30074711.html](https://paper.people.com.cn/rmr/pc/content/202505/22/content_30074711.html), 2024-09-20.
- [10] 任泽裕, 王振超, 柯尊旺, 等. 多模态数据融合综述[J]. 计算机工程与应用, 2021, 57(18): 49-64.
- [11] Huawei Technologies Co., Ltd. (2022) Huawei MindSpore AI Development Framework. In: Huawei Technologies Co., Ltd., Eds., *Artificial Intelligence Technology*, Springer, 137-162. [https://doi.org/10.1007/978-981-19-2879-6\\_5](https://doi.org/10.1007/978-981-19-2879-6_5)
- [12] Tong, Z., Du, N., Song, X. and Wang, X. (2021) Study on MindSpore Deep Learning Framework. 2021 17th International Conference on Computational Intelligence and Security (CIS), Chengdu, 19-22 November 2021, 183-186. <https://doi.org/10.1109/cis54983.2021.00046>
- [13] 王永威, 沈弢, 张圣宇, 等. 大小模型端云协同进化技术进展[J]. 中国图象图形学报, 2024, 29(6): 1510-1534.
- [14] 范建, 刘京, 武建双, 等. 基于人工智能, 大数据技术在智慧城市基层治理的应用与分析[J]. 人工智能与机器人

- 研究, 2025(14): 427-438.
- [15] Zhang, D., Yu, Y., Dong, J., Li, C., Su, D., Chu, C., *et al.* (2024) MM-LLMs: Recent Advances in MultiModal Large Language Models. *Findings of the Association for Computational Linguistics ACL 2024*, Bangkok, 11-16 August 2024, 12401-12430. <https://doi.org/10.18653/v1/2024.findings-acl.738>
- [16] Wu, J., Gan, W., Chen, Z., Wan, S. and Yu, P.S. (2023) Multimodal Large Language Models: A Survey. 2023 *IEEE International Conference on Big Data (BigData)*, Sorrento, 15-18 December 2023, 2247-2256. <https://doi.org/10.1109/bigdata59044.2023.10386743>
- [17] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image Is Worth  $16 \times 16$  Words: Transformers for Image Recognition at Scale. arXiv: 2010.11929.
- [18] Wang, Y., Huang, R., Song, S., *et al.* (2021) Not All Images Are Worth  $16 \times 16$  Words: Dynamic Transformers for Efficient Image Recognition. *Advances in Neural Information Processing Systems*, **34**, 11960-11973.
- [19] Masoudnia, S. and Ebrahimpour, R. (2012) Mixture of Experts: A Literature Survey. *Artificial Intelligence Review*, **42**, 275-293. <https://doi.org/10.1007/s10462-012-9338-y>
- [20] Xue, F., Zheng, Z., Fu, Y., *et al.* (2024) Openmoe: An Early Effort on Open Mixture-Of-Experts Language Models. arXiv: 2402.01739.
- [21] Wu, X., Huang, S. and Wei, F. (2024) Mixture of Lora Experts. arXiv: 2404.13628.
- [22] Mao, Y., Ge, Y., Fan, Y., Xu, W., Mi, Y., Hu, Z., *et al.* (2024) A Survey on Lora of Large Language Models. *Frontiers of Computer Science*, **19**, Article No. 197605. <https://doi.org/10.1007/s11704-024-40663-9>
- [23] Chen, J., Lin, H., Han, X. and Sun, L. (2024) Benchmarking Large Language Models in Retrieval-Augmented Generation. *Proceedings of the AAAI Conference on Artificial Intelligence*, **38**, 17754-17762. <https://doi.org/10.1609/aaai.v38i16.29728>
- [24] Mahajan, Y., Freestone, M., Aakur, S., *et al.* (2025) Revisiting Word Embeddings in the LLM Era. arXiv: 2502.19607.
- [25] Jiang, T., Song, M., Zhang, Z., *et al.* (2024) E5-v: Universal Embeddings with Multimodal Large Language Models. arXiv: 2407.12580.
- [26] Guo, W., Wang, J. and Wang, S. (2019) Deep Multimodal Representation Learning: A Survey. *IEEE Access*, **7**, 63373-63394. <https://doi.org/10.1109/access.2019.2916887>
- [27] Wang, J., Yi, X., Guo, R., Jin, H., Xu, P., Li, S., *et al.* (2021) Milvus: A Purpose-Built Vector Data Management System. *Proceedings of the 2021 International Conference on Management of Data*, 20-25 June 2021, 2614-2627. <https://doi.org/10.1145/3448016.3457550>