# LLM 自动化评估结果

以下是根据您的查询，我对提供的需求文档进行的完整分析。我将逐步执行您的指示，包括统计数据实体、合并功能需求、判断幻觉需求、评估完整引用和可测试性，并输出最终评估表。

### 步骤1: 统计文档中涉及的数据实体

数据实体是指系统内部需要管理的核心数据对象（如数据库表、数据模型中的实体），不包括外部数据源、接口或格式（如外部API或传感器）。基于功能需求文档，我提取了所有提及的数据实体，并确保不重复计数。

- \*\*提取的数据实体列表\*\*：从功能需求中识别，共9个 distinct 数据实体：

1. \*\*交通数据 (Traffic Data)\*\*：核心实体，涉及数据摄取、清洗、分析、过滤等。例如，INGEST-001、CLEAN系列、FILTER系列等。

2. \*\*数据源 (Data Source)\*\*：系统内部管理的来源信息，如MAPPING-001中“自动检测新数据源”。

3. \*\*数据模式/格式 (Data Schema/Format)\*\*：系统内部的数据结构，如MAPPING-001中“建议映射到现有数据模式”。

4. \*\*解决事件 (Resolved Event)\*\*：数据调和后的输出，如RECONCILE-001中“为每个已解决事件分配置信度分数”。

5. \*\*交通趋势 (Traffic Trend)\*\*：分析输出的派生数据，如ANALYSIS系列中“识别交通趋势”。

6. \*\*报告 (Report)\*\*：生成的输出对象，如REPORT-001中“生成可视化报告”。

7. \*\*数据请求 (Data Request)\*\*：跟踪的请求对象，如REQUEST系列中“支持数据请求跟踪系统”。

8. \*\*警报 (Alert)\*\*：操作状态数据，如DASHBOARD-003中“显示数据请求和警报的状态”。

9. \*\*问题 (Issue)\*\* 和 \*\*行动 (Action)\*\*：FLAG系列中“标记高影响问题并建议行动”；但注意，FLAG-001中的“减少库存水平”涉及非交通实体，可能部分无效（后续幻觉检查会处理）。

- \*\*排除项\*\*：外部接口中提到的数据（如天气API、公共安全数据库）不计入；约束部分未引入新数据实体。

- \*\*数据实体总数：9个\*\*。

### 步骤2: 合并功能需求并统计个数

根据您的规则，我对功能需求进行合并：

- \*\*规则1\*\*：对同一数据实体的相同功能点合并（如多个针对交通数据的分析需求合并为一个“交通分析”）。

- \*\*规则2\*\*：剔除系统通用设置类需求（如偏好或多语言支持），但文档中没有此类需求，因此无需剔除。

- \*\*合并过程\*\*：

- CLEAN-001/002/003（数据清洗）合并：所有针对交通数据的清洗任务（删除重复、标准化格式、标记无效字段）视为同一功能点。

- ANALYSIS-001/002/003（交通分析）合并：所有针对交通数据的趋势识别（高峰时段、瓶颈、异常）视为同一功能点。

- PREDICT-001/002/003（预测洞察）合并：所有提供预测性洞察（预测慢速移动产品、交通事件、建议措施）视为同一功能点，但“慢速移动产品”可能不协调。

- FLAG-001/002/003（问题标记）合并：所有标记问题并建议行动（减少库存水平、调整信号、安排维护）视为同一功能点，但“减少库存水平”可能不协调。

- FILTER-001/002/003（数据过滤）合并：所有针对交通数据的过滤（时间、星期、区域）视为同一功能点。

- DASHBOARD-001/002/003（仪表盘）合并：所有显示功能（实时数据、历史数据、请求状态）视为同一功能点。

- REQUEST-001/002（数据请求跟踪）合并：所有请求跟踪功能（请求表单、状态更新）视为同一功能点。

- 其他需求（如INGEST-001）未与其他合并，保持独立。

- \*\*合并后功能需求列表\*\*：共13个功能需求点：

1. \*\*数据摄取 (Data Ingestion)\*\*：合并自INGEST-001。

2. \*\*数据清洗 (Data Cleaning)\*\*：合并自CLEAN-001/002/003。

3. \*\*数据调和 (Data Reconciliation)\*\*：来自RECONCILE-001。

4. \*\*数据映射 (Data Mapping)\*\*：来自MAPPING-001。

5. \*\*交通分析 (Traffic Analysis)\*\*：合并自ANALYSIS-001/002/003。

6. \*\*预测洞察 (Predictive Insights)\*\*：合并自PREDICT-001/002/003。

7. \*\*问题标记 (Issue Flagging)\*\*：合并自FLAG-001/002/003。

8. \*\*报告 (Reporting)\*\*：来自REPORT-001。

9. \*\*数据过滤 (Data Filtering)\*\*：合并自FILTER-001/002/003。

10. \*\*仪表盘 (Dashboard)\*\*：合并自DASHBOARD-001/002/003。

11. \*\*数据请求跟踪 (Data Request Tracking)\*\*：合并自REQUEST-001/002。

12. \*\*数据共享 (Data Sharing)\*\*：来自SHARE-001。

13. \*\*数据可视化 (Data Visualization)\*\*：来自VISUALIZATION-001。

- \*\*合并后功能需求总数：13条\*\*。

### 步骤3: 判断幻觉需求

- \*\*评判标准\*\*：幻觉需求定义为引入功能操作，但涉及的数据实体无法与系统中任何现有实体关联，导致ER图不再是弱连通的（即存在孤立节点）。仅当需求核心部分脱节时才标记为幻觉；扩展描述不视为幻觉。

- \*\*分析\*\*：核心数据实体是交通数据（Traffic Data），其他实体应能通过路径与之连接。预测洞察中的“慢速移动产品”和问题标记中的“库存水平”不是交通相关实体，且未在系统中定义，导致孤立节点。

- \*\*一一列举判断\*\*：

- \*\*数据摄取\*\*：否。实体：交通数据、数据源。理由：实体在系统核心中定义（如摄取来源），与交通数据直接连接。

- \*\*数据清洗\*\*：否。实体：交通数据。理由：清洗操作直接处理交通数据，是核心功能的一部分。

- \*\*数据调和\*\*：否。实体：交通数据、解决事件。理由：解决事件派生自交通数据，通过数据调和路径连接。

- \*\*数据映射\*\*：否。实体：数据源、数据模式。理由：数据源在摄取中定义，模式与数据实体关联。

- \*\*交通分析\*\*：否。实体：交通数据、交通趋势。理由：趋势是交通数据的分析输出，直接连接。

- \*\*预测洞察\*\*：是。实体：预测洞察（包括“慢速移动产品”）。理由：引入“慢速移动产品”实体，未在交通系统中定义（如与库存管理相关），无法通过任何路径连接到交通数据或其他实体，导致ER图孤立节点。

- \*\*问题标记\*\*：是。实体：问题、行动（包括“减少库存水平”）。理由：引入“库存水平”实体，未在交通系统中定义（如与库存管理相关），无法连接到核心实体（如交通数据），导致孤立节点。

- \*\*报告\*\*：否。实体：报告、交通数据。理由：报告基于交通数据生成，通过严重程度路径连接。

- \*\*数据过滤\*\*：否。实体：交通数据。理由：过滤操作直接处理交通数据。

- \*\*仪表盘\*\*：否。实体：实时交通数据、历史数据、数据请求、警报。理由：所有实体在系统中定义（历史数据作为交通数据的子集），数据请求和警报与其他操作连接。

- \*\*数据请求跟踪\*\*：否。实体：数据请求。理由：请求实体在系统中管理，与数据共享或仪表盘连接。

- \*\*数据共享\*\*：否。实体：数据（部门间共享）。理由：共享数据基于交通数据或其他定义实体，通过治理策略连接。

- \*\*数据可视化\*\*：否。实体：实时数据。理由：数据来自外部但通过接口访问，实体可连接到系统核心（如仪表盘中的显示）。

- \*\*幻觉需求总数：2条\*\*（预测洞察和问题标记）。

- \*\*非幻觉需求总数：11条\*\*。

### 步骤4: 判断完整引用（针对非幻觉需求）

- \*\*评判标准\*\*：完整引用需求指功能描述中未引用未在需求规约说明书（包括功能需求和外部接口）中定义的功能、输入或输出。评判不严格，只要在文档中提及即可（如外部接口在约束5.3中定义）。

- \*\*分析\*\*：非幻觉需求的所有输入/输出均在文档中定义或引用。外部接口（如天气API）在约束5.3中描述。

- \*\*一一列举判断\*\*（仅非幻觉需求）：

- \*\*数据摄取\*\*：是。理由：输入（交通数据来源如摄像头、GPS）和输出（摄取数据）在需求中定义。

- \*\*数据清洗\*\*：是。理由：输入（原始交通数据）、输出（清洗后数据）和清洗任务在需求中定义。

- \*\*数据调和\*\*：是。理由：输入（冲突数据）、输出（解决事件和置信度分数）在需求中定义。

- \*\*数据映射\*\*：是。理由：输入（新数据源）、输出（映射建议）在需求中定义。

- \*\*交通分析\*\*：是。理由：输入（交通数据）、输出（交通趋势如瓶颈）在需求中定义。

- \*\*报告\*\*：是。理由：输入（交通数据）、输出（报告和颜色编码）在需求中定义。

- \*\*数据过滤\*\*：是。理由：输入（用户过滤条件）、输出（过滤后交通数据）在需求中定义。

- \*\*仪表盘\*\*：是。理由：输入（实时数据、历史数据、请求状态）和输出（显示）在需求中定义；外部接口如气象服务在约束5.3引用。

- \*\*数据请求跟踪\*\*：是。理由：输入（请求表单字段）、输出（状态更新）在需求中定义。

- \*\*数据共享\*\*：是。理由：输入/输出（共享数据）和治理策略在需求及约束5.1中定义。

- \*\*数据可视化\*\*：是。理由：输入（外部实时数据）、输出（可视化）在需求中定义；外部部门在约束5.3引用。

- \*\*非幻觉需求中完整引用总数：11条\*\*（全部为是）。

### 步骤5: 判断可测试性（针对非幻觉需求）

- \*\*评判标准\*\*：可测试需求指所有输入可以转化为所需输出，评判不严格；依赖外部接口或用户交互的可接受为可测试。

- \*\*分析\*\*：非幻觉需求均有明确输入和输出，可通过测试案例（如模拟数据或用户操作）验证转化。外部接口依赖不影响可测试性。

- \*\*一一列举判断\*\*（仅非幻觉需求）：

- \*\*数据摄取\*\*：是。理由：输入（如传感器数据）可通过测试数据验证输出（摄取状态）。

- \*\*数据清洗\*\*：是。理由：输入（含重复/无效数据）可转化为输出（清洗后数据），通过自动化测试验证。

- \*\*数据调和\*\*：是。理由：输入（冲突数据）可转化为输出（解决事件和置信度），通过模拟冲突测试。

- \*\*数据映射\*\*：是。理由：输入（新数据源）可转化为输出（映射建议），通过测试新源集成验证。

- \*\*交通分析\*\*：是。理由：输入（交通数据）可转化为输出（趋势识别），使用历史数据测试。

- \*\*报告\*\*：是。理由：输入（交通数据）可转化为输出（颜色编码报告），通过报告生成和验证测试。

- \*\*数据过滤\*\*：是。理由：输入（用户过滤条件）可转化为输出（过滤数据），通过应用条件测试。

- \*\*仪表盘\*\*：是。理由：输入（实时/历史数据）可转化为输出（显示），通过UI测试验证内容。

- \*\*数据请求跟踪\*\*：是。理由：输入（请求表单）可转化为输出（状态更新），通过提交请求测试。

- \*\*数据共享\*\*：是。理由：输入（数据）可转化为输出（共享数据），通过访问控制测试。

- \*\*数据可视化\*\*：是。理由：输入（外部数据）可转化为输出（可视化），通过API调用测试。

- \*\*非幻觉需求中可测试需求总数：11条\*\*（全部为是）。

### 步骤6: 输出完整评估表

根据以上分析，制作功能需求评估表。每条合并功能需求均被评估，无遗漏。幻觉需求标记为是，完整引用和可测试性因幻觉而标记为否。理由栏简要总结判断依据。

#### 功能需求评估表

| 功能需求 | 幻觉需求 | 完整引用 | 可行性 (可测试性) | 评判理由 |

|----------|----------|----------|-------------------|----------|

| 数据摄取 | 否 | 是 | 是 | 实体（交通数据、数据源）连接系统核心；输入/输出在需求中定义；可测试数据摄取过程。 |

| 数据清洗 | 否 | 是 | 是 | 实体（交通数据）连接系统核心；输入/输出在需求中定义；可测试清洗任务。 |

| 数据调和 | 否 | 是 | 是 | 实体（交通数据、解决事件）连接系统核心；输入/输出在需求中定义；可测试数据调和。 |

| 数据映射 | 否 | 是 | 是 | 实体（数据源、数据模式）连接系统核心；输入/输出在需求中定义；可测试映射功能。 |

| 交通分析 | 否 | 是 | 是 | 实体（交通数据、交通趋势）连接系统核心；输入/输出在需求中定义；可测试趋势分析。 |

| 预测洞察 | 是 | 否 | 否 | 引入“慢速移动产品”实体，无法与系统连接，导致ER图不连通；输入/输出部分未定义；不可测试（操作脱节）。 |

| 问题标记 | 是 | 否 | 否 | 引入“库存水平”实体，无法与系统连接，导致ER图不连通；输入/输出部分未定义；不可测试（操作不适用）。 |

| 报告 | 否 | 是 | 是 | 实体（报告、交通数据）连接系统核心；输入/输出在需求中定义；可测试报告生成。 |

| 数据过滤 | 否 | 是 | 是 | 实体（交通数据）连接系统核心；输入/输出在需求中定义；可测试过滤功能。 |

| 仪表盘 | 否 | 是 | 是 | 实体（实时数据、历史数据等）连接系统核心；输入/输出在需求中定义；可测试显示内容。 |

| 数据请求跟踪 | 否 | 是 | 是 | 实体（数据请求）连接系统核心；输入/输出在需求中定义；可测试跟踪系统。 |

| 数据共享 | 否 | 是 | 是 | 实体（共享数据）连接系统核心；输入/输出在需求中定义；可测试共享机制。 |

| 数据可视化 | 否 | 是 | 是 | 实体（外部数据）通过接口连接；输入/输出在需求中定义；可测试可视化功能。 |

### 总结

- \*\*整个文档涉及的数据实体有9个\*\*：交通数据、数据源、数据模式/格式、解决事件、交通趋势、报告、数据请求、警报、问题（问题标记中的“问题”实体，但部分操作无效）。

- \*\*经过相同功能点合并后，功能需求共13条\*\*。

- \*\*其中，幻觉需求有2条\*\*：预测洞察和问题标记。

- \*\*非幻觉需求有11条\*\*。

- \*\*非幻觉需求中完整引用的功能需求有11条\*\*。

- \*\*非幻觉需求中可测试的功能需求有11条\*\*。

\*\*评估说明\*\*：

- 我严格遵循了您的评判标准：幻觉需求基于ER图连通性（预测洞察和问题标记引入孤立实体）；完整引用和可测试性在非幻觉需求中全为“是”，因为文档定义清晰，但幻觉需求自动视为不完整且不可测试。

- 正常情况不可能所有需求完美：幻觉需求（如预测洞察中的“慢速移动产品”）是主要问题点，体现了需求不一致性（可能源于文档错误或域不匹配）。其他需求均符合标准。