# LLM 自动化评估结果

### 功能需求评估表

根据用户要求，我首先对功能需求进行合并统计，确保粒度一致。合并规则如下：

- 同一数据实体的相同功能点合并为一个功能需求（例如，CLEAN-001/002/003 合并为“数据清洗自动化”）。

- 剔除系统通用设置类需求（如偏好设置、多语言支持），但本需求文档中无此类需求。

- 避免过度合并：仅在需求高度相似或针对同一数据实体的同一操作时才合并。例如，PREDICT 和 FLAG 系列需求因涉及不同数据实体（如“slow-moving products”与交通事件）未合并。

- 统计后，功能需求共17条（合并后）。

数据实体识别：

- 仅统计系统内部管理的数据实体（基于功能需求和系统概述），排除外部接口（如数据源、格式）。

- 数据实体定义：系统内需要存储、处理的核心数据对象。从文档中提取的实体包括：

- 交通数据（来自传感器、摄像头等）。

- 天气数据（来自天气站）。

- 道路条件数据（如表面状况）。

- 事件数据（事故、类型、严重性）。

- 度量数据（拥堵水平、旅行时间等）。

- 数据质量数据（清洗和协调后的指标）。

- 警报数据（交通条件、天气事件等）。

- 请求数据（数据请求状态）。

- 总数据实体个数：8个（交通数据、天气数据、道路条件数据、事件数据、度量数据、数据质量数据、警报数据、请求数据）。缓存数据和存档数据被视为交通数据的子集，未单独计数。

幻觉需求判断标准：

- 幻觉需求：引入新操作，但涉及的数据实体无法与系统中任何现有实体关联（导致E-R图不再弱连接）。

- 非幻觉需求：操作涉及的数据实体能与现有实体关联。

- 注意：系统通用类需求已剔除，不参与统计。

完整引用判断标准（仅非幻觉需求）：

- 完整引用：需求描述的功能、输入或输出均在需求规约说明书（给定功能需求或约束）中有提及，无未定义元素。

- 宽松判断：如果内容在其他功能需求或外部接口中提到，即视为完整引用。

可测试性判断标准（仅非幻觉需求）：

- 可测试：需求的所有输入可转化为所需输出（例如，提供输入数据，可验证输出）。

- 宽松判断：即使依赖外部接口或用户交互，也视为可测试。

- 不可测试：需求模糊、主观或输入输出无法明确映射。

评估表（每条功能需求独立评判）：

| 功能需求 | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| 数据摄取（基于INGEST-001） | 否 | 是 | 是 | 涉及交通数据实体（系统核心），与现有实体关联；输入（来源如交通摄像头）和输出（摄取数据）在功能需求和系统概述中定义；输入（数据源）可测试输出（数据被摄取）。 |

| 数据清洗自动化（合并CLEAN-001/002/003） | 否 | 是 | 是 | 涉及数据质量数据实体（系统核心），与现有实体关联；输入（原始数据）和输出（清洗后数据）在需求中定义；输入（如重复条目）可测试输出（清洗结果）。 |

| 数据协调（RECONCILE-001） | 否 | 是 | 是 | 涉及数据质量数据实体（系统核心），与现有实体关联；输入（冲突数据）和输出（协调后数据及置信分数）在需求中定义；输入（冲突数据）可测试输出（协调结果）。 |

| 数据映射（MAPPING-001） | 否 | 是 | 是 | 涉及数据源模式，但映射目标为现有数据实体（如交通数据），可关联；输入（新数据源）和输出（映射建议）在需求中定义；输入（新源）可测试输出（建议映射）。 |

| 交通趋势分析（合并ANALYSIS-001/002/003） | 否 | 是 | 是 | 涉及交通数据实体（系统核心），与现有实体关联；输入（交通数据）和输出（趋势如高峰小时）在需求中定义；输入（数据）可测试输出（识别趋势）。 |

| 预测 slow-moving products（PREDICT-001） | 是 | 否 | 否 | 引入“slow-moving products”实体（如库存产品），但系统核心实体均为交通相关（无库存或产品实体），无法关联（E-R图断开）；输入输出未在规约中明确定义，引用不完整；因实体缺失，输入无法映射到输出，不可测试。 |

| 预测交通事件（PREDICT-002） | 否 | 是 | 是 | 涉及事件数据实体（系统核心），与现有实体关联；输入（交通数据）和输出（事件预测）在需求中定义；输入（数据）可测试输出（预测准确性）。 |

| 建议主动措施（PREDICT-003） | 否 | 是 | 是 | 不引入新实体，基于数据模式（系统核心）建议措施；输入（数据模式）和输出（建议）在需求中定义；输入（模式）可测试输出（建议生成）。 |

| 标记减少库存水平（FLAG-001） | 是 | 否 | 否 | 引入“stock levels”实体（库存水平），但系统核心实体为交通相关（无库存实体），无法关联（E-R图断开）；输入输出未在规约中明确定义，引用不完整；因实体缺失，输入无法映射到输出，不可测试。 |

| 标记调整交通信号（FLAG-002） | 否 | 是 | 是 | 涉及交通数据实体（系统核心，信号调整基于交通条件），与现有实体关联；输入（交通问题）和输出（调整建议）在需求中定义；输入（问题条件）可测试输出（建议）。 |

| 标记安排维护（FLAG-003） | 否 | 是 | 是 | 涉及道路条件或资产数据实体（系统核心），与现有实体关联；输入（维护需求）和输出（安排建议）在需求中定义；输入（需求）可测试输出（建议）。 |

| 报告交通严重性（REPORT-001） | 否 | 是 | 是 | 涉及度量数据实体（系统核心），与现有实体关联；输入（交通数据）和输出（颜色编码报告）在需求中定义；输入（数据）可测试输出（报告生成）。 |

| 数据过滤（合并FILTER-001/002/003） | 否 | 是 | 是 | 涉及交通数据实体（系统核心），与现有实体关联；输入（过滤参数如时间/区域）和输出（过滤后数据）在需求中定义；输入（参数）可测试输出（过滤结果）。 |

| 仪表板显示（合并DASHBOARD-001/002/003） | 否 | 是 | 是 | 涉及多种数据实体（交通、历史数据等），与现有实体关联；输入（实时/历史数据）和输出（显示状态）在需求中定义；输入（数据）可测试输出（显示内容）。 |

| 数据请求跟踪（合并REQUEST-001/002） | 否 | 是 | 是 | 涉及请求数据实体（系统核心），与现有实体关联；输入（请求表单）和输出（状态更新）在需求中定义；输入（请求）可测试输出（状态变化）。 |

| 数据共享（SHARE-001） | 否 | 是 | 是 | 涉及共享数据实体（系统核心），与现有实体关联；输入（部门数据）和输出（共享数据）在需求和约束中定义（合规政策）；输入（数据）可测试输出（共享成功）。 |

| 数据可视化（VISUALIZATION-001） | 否 | 是 | 是 | 涉及数据实体（系统核心），与现有实体关联；输入（实时数据）和输出（可视化）在需求中定义；输入（数据）可测试输出（可视化生成）。 |

### 总结

- \*\*整个文档涉及的数据实体有8个\*\*：交通数据、天气数据、道路条件数据、事件数据、度量数据、数据质量数据、警报数据、请求数据。

- \*\*经过相同功能点合并后，功能需求共17条\*\*。

- \*\*其中，幻觉需求有2条\*\*（PREDICT-001 和 FLAG-001），\*\*非幻觉需求有15条\*\*。

- \*\*非幻觉需求中，完整引用的功能需求有15条\*\*（所有非幻觉需求均完整引用）。

- \*\*非幻觉需求中，可测试的功能需求有15条\*\*（所有非幻觉需求均可测试）。

\*\*评审说明\*\*：

- 幻觉需求：PREDICT-001 和 FLAG-001 引入了未定义的数据实体（“slow-moving products”和“stock levels”），与系统核心交通实体脱节，导致E-R图不连通。

- 完整引用：非幻觉需求均无未定义元素，内容在其他功能需求或约束中有提及（如数据共享的合规在约束中定义）。

- 可测试性：非幻觉需求输入输出明确，可设计测试用例（如提供输入数据验证输出）。尽管预测和建议类需求可能涉及主观性，但在宽松标准下仍视为可测试。

- 严格评判：总需求中有2条幻觉需求，但完整引用和可测试性均无问题，这反映了需求文档的总体质量较高；幻觉需求主要源于需求描述中的不一致（如“products”在交通上下文中的不匹配）。