# LLM 自动化评估结果

### 功能需求评估表

根据您的要求，我对需求文档进行了详细评审。首先，基于功能需求，识别了系统需要管理的数据实体（不统计外部接口如外部数据源或格式）。然后，按照合并规则（同一数据实体的相同功能点合并，剔除系统通用设置类需求）对功能需求进行合并统计。接着，对每个功能需求逐一判断是否为幻觉需求、是否完整引用（仅针对非幻觉需求），以及是否可测试（仅针对非幻觉需求）。评判标准严格应用，确保不遗漏任何需求，并识别潜在问题。

#### 步骤1: 识别数据实体

从功能需求中提取的数据实体（系统内部管理的数据对象）：

- \*\*Scenario（场景）\*\*：来自SC系列需求，涉及创建、模拟和跟踪场景信息。

- \*\*Historical Data（历史数据）\*\*：来自SC-3，用于比较模拟结果和实际情况。

- \*\*Traffic Data（交通数据）\*\*：来自DA系列需求，包括原始和分析后的数据（如速度、流量变化）。

- \*\*Report（报告）\*\*：来自RG系列需求，包括生成、导出和类型管理。

- \*\*Collaboration Data（协作数据）\*\*：来自CO系列需求，涵盖评论、任务和工作空间版本信息。

数据实体总数：5个（Scenario、Historical Data、Traffic Data、Report、Collaboration Data）。

#### 步骤2: 合并功能需求

按照规则进行合并：

- \*\*规则1\*\*：同一数据实体的相同功能点合并。例如，SC-1、SC-2和SC-3均涉及Scenario实体，但功能点不同（创建、模拟、跟踪），因此不合并。类似地，其他需求均针对不同操作，无合并。

- \*\*规则2\*\*：剔除系统通用设置类需求（如偏好设置或多语言支持）。需求文档中无此类通用需求，因此无剔除。

- 合并后功能需求总数：14条（原始14个ID均保留）。

#### 步骤3: 判断幻觉需求

- \*\*评判标准\*\*：幻觉需求是引入功能操作时，涉及的数据实体无法与系统中任何现有实体关联，导致实体-关系（E-R）图不再弱连接（即存在孤立节点）。E-R图基于数据实体构建（Scenario、Historical Data、Traffic Data、Report、Collaboration Data），所有实体应通过路径连接（例如，Scenario 关联 Historical Data 和 Traffic Data，Traffic Data 用于生成 Report，Report 关联 Collaboration Data）。

- 非幻觉需求（即实体关联良好）不导致E-R图断裂。

- 所有需求均非幻觉，但严格评判下，部分需求边界模糊（如CO-1），但整体实体连接完整。

#### 步骤4: 判断完整引用需求（仅针对非幻觉需求）

- \*\*评判标准\*\*：需求完整引用是指该需求没有引用未在需求规约说明书（包括功能需求和外部描述）中定义的功能、输入或输出（包括数据库）。只要描述内容在其他需求或外部接口中有提及，即视为完整引用。

- 外部描述包含约束（如5.1-5.6），但不定义新功能实体，因此仅参考其提及内容。

#### 步骤5: 判断可测试需求（仅针对非幻觉需求）

- \*\*评判标准\*\*：可测试需求是所有输入可以达成所需输出的转化。依赖外部接口或用户交互可接受，不需特别严格（如输入输出可模拟）。

- 不可测试需求通常输入输出不明确或无法验证转化。

最终评估表如下。每个功能需求均被评判，评判理由基于需求描述和标准。

### 功能需求评估表

| 功能需求 ID | 功能需求描述 | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| SC-1 | Scenario Creation: Allow users to create custom traffic management scenarios including lane closures, signal timing changes, and detour routes. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Scenario实体，与Historical Data（SC-3）和Traffic Data（DA-1）关联，E-R图连接完整。<br>\*\*完整引用：是\*\* 输入（用户定义参数）和输出（Scenario实体）在需求中定义；外部接口（如道路数据）在系统概述提及。<br>\*\*可测试：是\*\* 输入场景参数可验证输出（保存的场景文件或状态）。 |

| SC-2 | Scenario Simulation: Simulate the impact of defined traffic scenarios on traffic congestion and travel times. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Scenario和Traffic Data实体，通过模拟输出（如拥堵指标）关联Historical Data（SC-3），E-R图连接。<br>\*\*完整引用：是\*\* 输入（Scenario）和输出（模拟结果）在SC-3和系统概述提及（交通指标）；外部依赖（如数据源）已定义。<br>\*\*可测试：是\*\* 输入场景可运行模拟，输出拥堵水平和旅行时间可测量。 |

| SC-3 | Historical Data Tracking: Maintain historical data tracking to compare simulated results with real-world outcomes. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Historical Data实体，与Scenario（SC-1/2）和Traffic Data（DA-1）关联，E-R图核心部分。<br>\*\*完整引用：是\*\* 输入（模拟结果和真实数据）和输出（比较结果）在需求中明确；数据源在系统概述定义。<br>\*\*可测试：是\*\* 输入数据集可比较输出差异（如偏差率）。 |

| DA-1 | Traffic Data Analysis: Analyze traffic data to provide insights such as average traffic speed, volume changes over time, and predictive modeling of traffic strategies. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Traffic Data实体，与Report（RG系列）和Historical Data（SC-3）关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（交通数据）在系统概述提及；输出（insights）在需求描述；预测建模在DA-1自身定义。<br>\*\*可测试：是\*\* 输入交通数据可生成输出指标（如平均速度），用户交互或算法可验证。 |

| DA-2 | Data Visualization: Render built-in visualizations, including heatmaps, line graphs, bar charts, and time-lapse comparisons. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Traffic Data实体（可视化基于此），与Report（RG系列）关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（分析数据）在DA-1提及；输出（可视化类型）在需求明确；无未定义引用。<br>\*\*可测试：是\*\* 输入数据可生成输出图形，渲染结果可检查。 |

| DA-3 | Data Filtering and Drill-Down: Allow users to filter and drill down into specific data points for in-depth analysis by technical teams. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Traffic Data实体（过滤操作），与DA-1分析关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（数据点）和输出（过滤视图）在需求描述；依赖项在DA-1和系统概述提及。<br>\*\*可测试：是\*\* 输入过滤条件可验证输出数据子集。 |

| RG-1 | Automated Report Generation: Generate automated reports using customizable templates. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Report实体，与Traffic Data（DA系列）和Collaboration Data（CO系列）关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（模板和数据）在RG-2/RG-4提及；输出（报告）在需求定义；无未定义引用。<br>\*\*可测试：是\*\* 输入模板和数据可生成输出报告文件。 |

| RG-2 | Report Export: Support exporting reports in multiple formats, including PDF, Excel, PowerPoint, and Google Slides. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Report实体（导出操作），与RG-1生成关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（报告）和输出（格式文件）在需求明确；格式列表已定义。<br>\*\*可测试：是\*\* 输入报告可导出为指定格式文件并验证。 |

| RG-3 | Recurring Report Exports: Allow recurring exports of reports to reduce manual effort. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Report实体（定期操作），与RG-2导出关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（报告和调度参数）在需求描述；依赖RG-1/RG-2。<br>\*\*可测试：是\*\* 输入调度设置可自动化输出导出并验证频率。 |

| RG-4 | Summary and Detailed Reports: Produce summary reports for leadership and detailed reports for technical teams. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Report实体（类型管理），与RG-1生成关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（数据）和输出（报告类型）在需求明确；无未定义引用。<br>\*\*可测试：是\*\* 输入数据可生成不同报告类型并检查内容差异。 |

| CO-1 | Collaboration Integration: Integrate with collaboration platforms such as WeCom, Feishu, and Microsoft Teams. | 否 | 否 | 是 | \*\*幻觉：否\*\* 涉及Collaboration Data实体（集成操作），与Report（RG系列）关联（如共享报告），E-R图连接。<br>\*\*完整引用：否\*\* 输入（平台接口）和输出（集成状态）未在需求中具体定义（仅列出平台名称，无API细节）；外部接口约束（5.6）提及加密，但不定义功能细节，因此引用不完整。<br>\*\*可测试：是\*\* 输入平台凭据可测试输出连接状态（依赖外部，但可模拟）。 |

| CO-2 | Real-Time Commenting and Highlighting: Allow real-time commenting and highlighting within shared reports. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Collaboration Data实体（评论），与Report（RG系列）关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（用户输入）和输出（评论在报告中）在需求描述；报告实体在RG系列定义。<br>\*\*可测试：是\*\* 输入评论内容可实时输出在报告中验证。 |

| CO-3 | Task Assignment and Change Tracking: Support task assignment and change tracking for collaborative workflows. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Collaboration Data实体（任务），与工作流关联，E-R图连接（例如链接到Scenario或Report）。<br>\*\*完整引用：是\*\* 输入（任务细节）和输出（跟踪状态）在需求明确；协作上下文在CO-1提及。<br>\*\*可测试：是\*\* 输入任务分配可输出跟踪日志并验证。 |

| CO-4 | Shared Workspace with Version Control: Provide a shared workspace with version control for team collaboration. | 否 | 是 | 是 | \*\*幻觉：否\*\* 涉及Collaboration Data实体（工作空间），与报告和任务关联，E-R图连接。<br>\*\*完整引用：是\*\* 输入（工作空间操作）和输出（版本历史）在需求描述；无未定义引用。<br>\*\*可测试：是\*\* 输入修改可输出版本记录并检查历史。 |

#### 总结

- \*\*整个文档涉及的数据实体有 5 个\*\*：Scenario、Historical Data、Traffic Data、Report、Collaboration Data。

- \*\*经过相同功能点合并后，功能需求共 14 条\*\*：无合并（需求均针对不同数据实体或操作），无通用设置需求剔除。

- \*\*其中，幻觉需求有 0 条\*\*：所有需求均关联现有数据实体，E-R图弱连接（无孤立节点）。

- \*\*非幻觉需求有 14 条\*\*：幻觉需求为0，因此全部为非幻觉。

- \*\*非幻觉需求中完整引用的功能需求有 13 条\*\*：仅CO-1因集成细节未定义，不完整引用。

- \*\*非幻觉需求中可测试的功能需求有 14 条\*\*：所有需求输入输出可转化，依赖外部或用户交互可接受。

本评估严格遵循您的标准。幻觉需求为0，是因为所有功能操作均绑定到核心数据实体，E-R图完整；完整引用问题仅出现在CO-1，因外部平台集成细节未具体化；可测试性无问题，因需求输入输出基本明确。如果需进一步调整评判严格度，请提供反馈！