# LLM 自动化评估结果

### 功能需求评估表

根据您的详细要求，我对需求文档进行了系统分析。以下是我的评估过程：

#### 步骤概述：

1. \*\*识别数据实体\*\*：从功能需求中提取系统内部管理的核心数据对象（不包括外部接口、外部数据源或格式）。最终识别出 \*\*11个数据实体\*\*：User（用户）、Route（路线）、Disruption（中断）、Alert（警报）、Alternative Route（替代路线）、Real-time Data（实时数据）、Preference（用户偏好）、Review（评论）、Historical Reliability Data（历史可靠性数据）、Crowd Level（拥挤程度）、Walking Direction（步行方向）。

2. \*\*合并功能需求\*\*：应用您的规则合并功能需求。

- \*\*规则1\*\*：合并同一数据实体的相同功能点（例如，FR-1-001 和 FR-1-002 都涉及中断警报，合并为1个需求）。

- \*\*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），但本文档中无此类需求，因此无剔除。

- 合并后，功能需求总数：\*\*10条\*\*（避免过度合并，仅当需求高度相似或针对同一实体操作时才合并）。

3. \*\*判断幻觉需求\*\*：基于“引入功能操作且数据实体无法与现有实体关联，导致E-R图非弱连接”的标准评估。合并后需求中无幻觉需求（所有实体均能关联）。

4. \*\*判断完整引用\*\*：对于非幻觉需求，判断是否完整引用（即所有功能、输入或输出均在需求规约中定义；外部接口或约束提到的可接受）。评估稍严格，但允许引用约束部分。

5. \*\*判断可测试性\*\*：对于非幻觉需求，判断是否可测试（即所有输入可转化为输出；依赖外部接口或用户交互的可视为可测试）。评估稍严格。

6. \*\*制作评估表\*\*：每条需求逐一评估，包括幻觉、完整引用（仅非幻觉时）、可测试性（仅非幻觉时）及理由。

#### 功能需求评估表

| 功能需求 (基于合并后需求) | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| R1: 系统提供实时中断警报包括原因和持续时间（合并FR-1-001和FR-1-002） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Alert和Disruption，可关联到User和Route；\*\*完整引用\*\*: 输入（中断事件）、输出（警报内容）在需求中定义；\*\*可测试\*\*: 输入中断检测可触发输出警报（可模拟事件验证）。 |

| R2: 系统在检测到中断时重新计算并更新路线（合并FR-2-001和FR-2-002） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Route和Disruption，可关联；\*\*完整引用\*\*: 输入（中断）、输出（更新路线）在需求中定义；\*\*可测试\*\*: 输入中断可触发重新计算算法（输出可验证）。 |

| R3: 系统建议上下文感知的替代路线基于实时条件和用户偏好（合并FR-3-001和FR-3-002） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Alternative Route和Preference，可关联到User和Route；\*\*完整引用\*\*: 输入（用户位置、偏好）、输出（替代路线）在需求中定义；\*\*可测试\*\*: 输入位置和偏好可生成输出路线建议（可测试不同场景）。 |

| R4: 系统为每个替代路线提供权衡解释包括时间、换乘次数和步行距离（基于FR-4-001） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Alternative Route，可关联到Route；\*\*完整引用\*\*: 输入（路线数据）、输出（解释）在需求中明确定义；\*\*可测试\*\*: 输入路线属性可转化为输出解释（可验证内容准确性）。 |

| R5: 系统实时集成本地交通数据源以提供最新信息包括替换服务（合并FR-5-001和FR-5-002） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Real-time Data，可关联到Route；\*\*完整引用\*\*: 输入/输出在需求中定义，外部数据源在约束中提到（SRL-3.5等）；\*\*可测试\*\*: 输入外部API调用可获取输出最新数据（可测试集成点）。 |

| R6: 系统在高峰时段显示公共交通拥挤指标（基于FR-6-001） | 否 | 否 | 否 | \*\*幻觉\*\*: 涉及实体Crowd Level，可关联到Public Transport选项；\*\*完整引用\*\*: 输出（指标）定义，但输入（如何获取拥挤数据）未在文档中定义（仅提及"高峰时段"，但数据源和格式未说明）；\*\*可测试\*\*: 输入未定义（如数据来源或算法），无法确保输入到输出的转化（依赖未指定源）。 |

| R7: 系统提供带视觉支持的基于地标的步行方向（合并FR-7-001和FR-7-002） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Walking Direction，可关联到Route；\*\*完整引用\*\*: 输入（用户位置）、输出（方向）在需求中定义；\*\*可测试\*\*: 输入位置可生成输出方向和可视化（可测试导航功能）。 |

| R8: 系统允许用户对替代路线进行可选评分和评论（合并FR-8-001和FR-8-002） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Review，可关联到User和Alternative Route；\*\*完整引用\*\*: 输入（用户反馈）、输出（评分存储）在需求中定义；\*\*可测试\*\*: 输入用户操作可触发输出评论存储（可测试UI和逻辑）。 |

| R9: 系统显示替代路线的历史可靠性数据（基于FR-9-001） | 否 | 否 | 否 | \*\*幻觉\*\*: 涉及实体Historical Reliability Data，可关联到Alternative Route；\*\*完整引用\*\*: 输出（数据显示）定义，但输入（历史数据来源或格式）未在文档中定义；\*\*可测试\*\*: 输入未定义（如数据获取方式），无法达成输出转化（依赖未知源）。 |

| R10: 系统对路线建议进行排名并突出首选（合并FR-10-001和FR-10-002） | 否 | 是 | 是 | \*\*幻觉\*\*: 涉及实体Alternative Route，可关联到Preference；\*\*完整引用\*\*: 输入（实时条件、偏好）、输出（排名列表）在需求中定义；\*\*可测试\*\*: 输入条件可生成输出排名（可测试算法逻辑）。 |

#### 总结

- \*\*整个文档涉及的数据实体有11个\*\*：User（用户）、Route（路线）、Disruption（中断）、Alert（警报）、Alternative Route（替代路线）、Real-time Data（实时数据）、Preference（偏好）、Review（评论）、Historical Reliability Data（历史可靠性数据）、Crowd Level（拥挤程度）、Walking Direction（步行方向）。

- \*\*经过相同功能点合并后，功能需求共10条\*\*。

- \*\*其中，幻觉需求有0条\*\*：所有需求涉及的数据实体均能与核心系统实体（如User、Route）关联，未导致E-R图非弱连接。

- \*\*非幻觉需求有10条\*\*（所有需求均为非幻觉）。

- \*\*非幻觉需求中完整引用的功能需求有8条\*\*：R6和R9不完整，因输入（数据源）未明确定义。

- \*\*非幻觉需求中可测试的功能需求有8条\*\*：R6和R9不可测试，因输入未定义无法达成输出转化。

\*\*评估说明\*\*：

- \*\*幻觉需求判断\*\*：所有需求均基于系统核心功能（导航和中断处理），实体如Crowd Level和Historical Reliability Data虽为扩展，但能通过Route或User关联（例如，Crowd Level关联公共交通选项，Historical Data关联替代路线），未引入孤立实体。

- \*\*完整引用判断\*\*：R6和R9被标记为“否”，因文档中未定义数据来源（如如何获取拥挤指标或历史数据），仅描述输出行为，违反了“输入必须定义”的标准。

- \*\*可测试性判断\*\*：R6和R9被标记为“否”，因输入未定义（如数据获取机制），无法进行可靠测试；其他需求虽依赖外部接口（如R5），但可模拟或测试。

- \*\*严格性\*\*：评估稍严格（如R6和R9问题），确保不完美需求被识别。所有需求均有评估，无遗漏。