# LLM 自动化评估结果

### 功能需求评估表

根据用户要求，我对功能需求进行了合并统计，以减少粒度并确保清晰表达核心功能和目的。合并规则如下：

- 对同一数据实体的相同功能点合并为一个功能需求（例如，所有拥挤度相关需求合并为“显示拥挤度信息”）。

- 剔除系统通用设置类需求（如多语言支持），但本需求文档中无此类需求。

- 避免过度合并：仅当需求非常相近或针对同一数据实体的同一操作时才合并。

经过合并后，功能需求共5条。以下是对每条功能需求的详细评估：

- \*\*幻觉需求评估标准\*\*：需求引入一个功能操作，涉及的数据实体无法与系统中任何现有实体关联（导致ER图不再弱连接）。若需求仅扩展描述现有实体，则不视为幻觉。

- \*\*完整引用评估标准\*\*：需求未引用未在需求文档（包括功能需求和外部接口）中定义的功能、输入或输出。评估不严格，只要文档中提及即可。

- \*\*可测试性评估标准\*\*：需求的所有输入可转化为所需输出（不严格评估，依赖外部接口或用户交互的需求可视为可测试）。

- 评估基于需求文档内容，并稍严格评判。

| 功能需求 (合并后描述) | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| 提供实时运输延误更新（来自 NAV-RT-001） | 是 | - | - | 该需求涉及数据实体“延误信息”，但该实体未与任何其他实体（如路线或用户）关联。在ER图中，延误信息节点孤立，导致图不连通（非弱连接）。 |

| 基于实时交通和施工条件建议替代路线，并标记原因及显示影响（来自 NAV-ALT-002、NAV-ALT-010、NAV-ALT-011） | 否 | 是 | 是 | 数据实体“交通信息”和“路线信息”已定义（交通信息包括施工，路线信息包括替代路线）。输入（实时交通数据）、输出（建议路线）在文档中引用（如外部接口 C-INT-001）。可通过模拟交通数据测试输出。 |

| 显示拥挤度信息，包括颜色编码、数值估计和文本描述（来自 NAV-CD-003、NAV-CD-004、NAV-CD-005、NAV-CD-006） | 否 | 是 | 是 | 数据实体“拥挤度信息”已定义，并通过用户偏好（如 NAV-USER-012）与“路线信息”关联。输入（拥挤度数据）、输出（视觉显示）在文档中引用（如 NAV-CD-003 和外部接口 C-INT-001）。可通过提供密度数据测试显示输出。 |

| 提供个性化路线建议，包括跟踪用户行为、学习偏好、允许反馈及设置时间偏好（来自 NAV-USER-007、NAV-USER-008、NAV-USER-009、NAV-USER-012、NAV-USER-014、NAV-USER-015、NAV-USER-016） | 否 | 是 | 是 | 数据实体“用户档案”和“路线信息”已定义并关联（用户档案包括行为、偏好）。输入（用户行为）、输出（建议路线）在文档中完整引用（如 NAV-USER-007 和 NAV-USER-008）。可测试（如使用历史数据验证建议逻辑，依赖用户交互）。 |

| 定位并显示附近自行车/滑板车站作为替代方案（来自 NAV-ALT-013） | 是 | - | - | 该需求涉及数据实体“车站信息”，但该实体未与任何其他实体（如路线或用户）关联。在ER图中，车站信息节点孤立，导致图不连通（非弱连接）。 |

### 总结

- \*\*整个文档涉及的数据实体个数\*\*：6个。数据实体是从功能需求中提取的系统内管理核心对象（不包括外部接口如数据源）。实体列表：

1. 延误信息（来自 NAV-RT-001）

2. 交通信息（包括施工，来自 NAV-ALT-002）

3. 拥挤度信息（来自 NAV-CD-003 至 006）

4. 路线信息（包括主路线和替代路线，来自 NAV-ALT-002、NAV-ALT-010、NAV-ALT-011、NAV-USER-008、NAV-USER-012）

5. 用户档案（包括行为、偏好和反馈，来自 NAV-USER-007、NAV-USER-008、NAV-USER-009、NAV-USER-012、NAV-USER-014、NAV-USER-015、NAV-USER-016）

6. 车站信息（自行车/滑板车站，来自 NAV-ALT-013）

- \*\*合并后功能需求个数\*\*：5条（基于合并规则，原始16条需求合并为5条）。

- \*\*幻觉需求个数\*\*：2条（RT\_Delay 和 ALT\_BikeStations），占合并后需求的40%。

- \*\*非幻觉需求个数\*\*：3条（ALT\_Suggest、CD\_Display、User\_Personalize），占合并后需求的60%。

- \*\*非幻觉需求中完整引用的功能需求个数\*\*：3条（所有非幻觉需求均完整引用），占非幻觉需求的100%。

- \*\*非幻觉需求中可测试的功能需求个数\*\*：3条（所有非幻觉需求均可测试），占非幻觉需求的100%。

\*\*评估说明\*\*：

- 幻觉需求识别严格：RT\_Delay 和 ALT\_BikeStations 引入的实体（延误信息和车站信息）孤立，系统核心架构（如路线和用户）未引用它们，导致ER图不连通。

- 完整引用评估宽松：非幻觉需求的功能、输入和输出均在文档中描述或通过外部接口（如 C-INT-001）引用。

- 可测试性评估合理：非幻觉需求均有明确输入和输出，可设计测试用例（如模拟数据或用户交互）。

- 总体稍严格：尽管非幻觉需求均通过评估，但幻觉需求比例较高，反映部分需求与系统核心脱节。