# LLM 自动化评估结果

### 功能需求评估表

以下是对City Mapper需求文档的完整评估。首先，我已根据用户要求，对功能需求进行了合并统计（基于同一数据实体的相同功能点），并剔除了系统通用设置类需求（如多语言支持或系统偏好设置，但用户偏好管理属于核心功能，未剔除）。接着，我逐步评估了每个合并后的功能需求是否为幻觉需求、是否完整引用，以及是否可测试。评估标准基于用户提供的定义，并稍严格评判，以确保合理性。

#### 关键步骤总结：

- \*\*数据实体识别\*\*：从功能需求中提取系统内需要管理的数据实体（不包括外部接口如外部数据源）。合并相关实体后，共识别出10个数据实体：Route Information、Traffic Conditions、Crowd Density Data、User Behavior Data、Personalized Model、User Feedback、Alternative Route Details、Preference-based Route Suggestion、Alternative Transport Stations、User Preferences。

- \*\*功能需求合并\*\*：原始16个功能需求被合并为8个合并需求（基于相同数据实体的相同操作）。合并规则：仅当需求针对同一数据实体的同一功能点时合并（例如，所有拥挤密度相关需求合并为一个）。系统通用设置类需求（如多语言支持）在文档中未出现，因此未剔除任何需求。

- \*\*幻觉需求判断\*\*：定义：需求引入的功能操作涉及的数据实体无法与系统现有实体关联，导致E-R图非弱连接（即存在孤立节点）。所有需求的数据实体均与核心实体（如Route、User）连接，因此无幻觉需求。

- \*\*完整引用判断\*\*：定义：需求不引用未在需求文档（包括功能需求和约束）中定义的功能、输入或输出。文档中所有输入/输出均有相关约束（如C-INT系列）或功能需求提及，因此均完整引用。

- \*\*可测试性判断\*\*：定义：所有输入可转化为所需输出（评估不严格，依赖外部接口或用户交互也视为可测试）。所有需求均有明确输入输出，因此均可测试。但稍严格评判后，仍无问题。

- \*\*评估细节\*\*：每个合并需求对应原始需求ID，并进行独立评判。表格中“功能需求”列为合并后的描述。

| 功能需求 | 幻觉需求 | 完整引用 | 可测试性 | 评判理由 |

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

| \*\*FR1: 提供实时公共交通延误更新\*\*（基于NAV-RT-001） | 否 | 是 | 是 | 数据实体“Route Information”与系统核心（路线导航）关联；输入源在约束C-INT-001定义（官方数据）；输出为延误更新，可通过API测试。 |

| \*\*FR2: 检测实时交通和施工条件并建议替代路线\*\*（基于NAV-ALT-002） | 否 | 是 | 是 | 数据实体“Traffic Conditions”和“Alternative Route Details”与路线实体连接；输入在C-INT-001定义；输出为建议路线，依赖外部接口但可测试。 |

| \*\*FR3: 提供拥挤密度信息（包括颜色编码、数值估计和文本描述）\*\*（合并NAV-CD-003,004,005,006） | 否 | 是 | 是 | 数据实体“Crowd Density Data”与公共交通选项关联；输入源在约束C-INT-001定义；输出显示密度信息，可模拟数据测试颜色、数值和文本输出。 |

| \*\*FR4: 管理用户个性化模型（包括跟踪用户行为、预测偏好路线、收集反馈）\*\*（合并NAV-USER-007,008,009） | 否 | 是 | 是 | 数据实体“User Behavior Data”、“Personalized Model”和“User Feedback”与用户偏好实体连接；输入输出在C-REG系列约束定义（数据保护），可测试行为跟踪和反馈机制。 |

| \*\*FR5: 提供替代路线详情（包括原因标签和旅行时间/成本影响）\*\*（合并NAV-ALT-010,011） | 否 | 是 | 是 | 数据实体“Alternative Route Details”与路线实体直接连接；输入在FR2相关约束中定义；输出原因和影响，可通过场景测试。 |

| \*\*FR6: 建议与用户偏好一致的路线（如较少拥挤、更安静或风景路线）\*\*（基于NAV-USER-012） | 否 | 是 | 是 | 数据实体“Preference-based Route Suggestion”与用户偏好和路线实体关联；输入偏好来自FR8；输出路线建议，依赖个性化模型但可测试。 |

| \*\*FR7: 显示附近自行车或滑板车站点作为公共交通替代\*\*（基于NAV-ALT-013） | 否 | 是 | 是 | 数据实体“Alternative Transport Stations”与替代路线实体连接；输入源在C-INT-001定义；输出站点信息，可通过位置数据测试。 |

| \*\*FR8: 管理用户偏好（包括设置时间偏好、记住偏好、学习偏好）\*\*（合并NAV-USER-014,015,016） | 否 | 是 | 是 | 数据实体“User Preferences”与用户核心实体连接；输入输出在C-REG约束定义；可测试偏好设置和记忆功能（学习过程虽需时间，但定义不严格）。 |

#### 总结

- \*\*整个文档涉及的数据实体有10个\*\*：Route Information（路线信息）、Traffic Conditions（交通条件）、Crowd Density Data（拥挤密度数据）、User Behavior Data（用户行为数据）、Personalized Model（个性化模型）、User Feedback（用户反馈）、Alternative Route Details（替代路线详情）、Preference-based Route Suggestion（偏好路线建议）、Alternative Transport Stations（替代交通站点信息）、User Preferences（用户偏好）。

- \*\*经过相同功能点合并后，功能需求共8条\*\*（原始16条需求通过合并减少）。

- \*\*其中，幻觉需求有0条\*\*：所有需求的数据实体均与系统核心实体（如用户、路线）连接，E-R图保持弱连接，无孤立实体。

- \*\*非幻觉需求有8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\*\*：所有需求的输入/输出均在文档中的功能需求或约束（如C-INT、C-REG系列）有明确定义，无未引用元素。

- \*\*非幻觉需求中可测试的功能需求有8条\*\*：所有需求均有明确输入（如实时数据、用户行为）和输出（如路线建议、偏好设置），可转化为测试用例，依赖外部接口（如API）或用户交互不影响可测试性。

\*\*附加说明\*\*：

- \*\*合并细节\*\*：合并基于数据实体和功能操作的同质性。例如：

- NAV-CD-003至006均针对“Crowd Density Data”的显示操作，因此合并为FR3。

- NAV-USER-007、008、009均针对“Personalized Model”的管理操作，因此合并为FR4。

- 系统通用设置类需求（如多语言支持）在文档中未出现，因此未剔除任何需求。

- \*\*严格评判\*\*：尽管用户指示稍严格评判，但需求文档较完整，所有需求均符合非幻觉、完整引用和可测试标准。潜在弱点（如FR8的学习偏好需时间测试）在定义下不影响可测试性。约束部分（如C-INT-001）确保了外部数据的可集成性。