# LLM 自动化评估结果

### 功能需求评估表

根据用户提供的需求文档，我将按照指定步骤进行评估。首先，我识别了数据实体，然后合并了功能需求（确保粒度相同）。接下来，我对每个合并后的功能点逐一评估了是否为幻觉需求、是否完整引用以及是否可测试。评估基于以下标准：

- \*\*幻觉需求\*\*：如果需求引入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（即存在孤立节点），则判定为幻觉需求。否则，非幻觉。

- \*\*完整引用\*\*：如果需求描述的功能、输入或输出在文档中的其他功能需求或外部接口中有提及或定义，则判定为完整引用（评判不严格）。

- \*\*可测试性\*\*：如果需求的输入可以达成所需输出的转化（例如，可通过测试用例验证），则判定为可测试（评判不严格；依赖外部接口或用户交互的也视为可测试）。

文档中的功能需求经合并后共有7条。合并规则如下：

- 同一数据实体的相同功能点合并（例如，LOG-001至LOG-004均针对Complaint实体的记录操作，合并为“Complaint Logging”）。

- 剔除系统通用设置类需求（如偏好设置或多语言支持），但文档中无此类需求，故未剔除。

- 合并时避免过度：仅在功能点非常相似或针对同一实体的同一操作时才合并（例如，ROUTE-001至ROUTE-004均涉及投诉路由，合并为“Complaint Routing”）。

数据实体识别仅从功能需求中提取系统内部管理的实体（外部接口提到的外部数据源或格式不统计）。识别出的数据实体有5个：\*\*Complaint\*\*（投诉记录）、\*\*Notification\*\*（通知）、\*\*AuditTrail\*\*（审计跟踪）、\*\*Department\*\*（部门）和\*\*Communication\*\*（通信）。这些实体在E-R图中以Complaint为核心，相互连接（弱连接）。

以下为完整评估表。每个功能需求基于合并后的点进行评估（原始ID在理由中说明）。评估稍严格，确保不遗漏任何点。

| 功能需求 | 幻觉需求 | 完整引用 | 可行性（可测试） | 评判理由 |

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

| Complaint Logging (合并LOG-001至004) | 否 | 是 | 是 | \*\*幻觉理由：\*\* 涉及数据实体Complaint，该实体在系统中存在，并与其他实体连接（如通过分类、路由），E-R图保持弱连接。\*\*完整引用理由：\*\* 输入（如contact details, complaint type）和输出（记录投诉）在LOG系列需求中定义，外部接口（如CST-IF-002）提到安全传输，支持输入输出定义。\*\*可测试理由：\*\* 输入（投诉数据）可模拟（如通过tablet输入），输出（记录投诉）可验证（检查记录是否完整），依赖用户交互但可测试。 |

| Complaint Classification (合并CLASS-001至003) | 否 | 是 | 是 | \*\*幻觉理由：\*\* 涉及数据实体Complaint，该实体存在并连接其他实体（如通过通知），E-R图无孤立节点。\*\*完整引用理由：\*\* 输入（关键词、投诉类型）和输出（分类结果）在CLASS系列需求中定义，CLASS-002提到“key phrases”，输出在NOTIFY-001中被引用（高风险分类触发通知）。\*\*可测试理由：\*\* 输入（投诉样本）可提供，输出（分类结果）可检查（如验证自动分类逻辑），并可测试用户确认（CLASS-003）。 |

| Notification Management (合并NOTIFY-001至003) | 否 | 否 | 是 | \*\*幻觉理由：\*\* 涉及数据实体Notification，该实体在系统中存在（基于Complaint实体），并连接其他实体（如通过路由或通知），E-R图保持连接。\*\*完整引用理由：\*\* \*\*不完整引用\*\*，因为输入（“high-risk”状态）在CLASS-001定义，但输出（通知形式如“dashboard alerts or direct messages”）未在其他地方定义详细格式；NOTIFY-003的“ensure urgency”未量化，引用不足。\*\*可测试理由：\*\* 输入（高风险投诉）可触发，输出（发送通知）可验证（如检查是否在24小时内升级或通过指定渠道送达），尽管依赖外部系统但可测试。 |

| Complaint Routing (合并ROUTE-001至004) | 否 | 是 | 是 | \*\*幻觉理由：\*\* 涉及数据实体Complaint和Department，均存在（Department在ROUTE中引入），并连接其他实体（如通过分类），E-R图无孤立。\*\*完整引用理由：\*\* 输入（分类、关键词、位置）在ROUTE和CLASS系列中定义，输出（路由到部门）在ROUTE中详细描述，外部接口（如CST-IF-001集成CRM）支持数据交换。\*\*可测试理由：\*\* 输入（投诉属性）可设置，输出（路由分配）可检查（如验证负载均衡或优先级处理），可模拟部门负载测试。 |

| Smart Suggestions (SUPP-001) | 否 | 是 | 是 | \*\*幻觉理由：\*\* 涉及数据实体Complaint和Department，均存在并连接系统核心，E-R图无孤立。\*\*完整引用理由：\*\* 输入（关键词）和输出（建议类型和部门）在SUPP-001定义，部门在ROUTE中引用，输入输出在功能需求中覆盖。\*\*可测试理由：\*\* 输入（关键词）可输入，输出（建议）可验证（如检查建议准确性），用户交互可测试。 |

| Maintain Audit Trail (SUPP-002) | 否 | 否 | 是 | \*\*幻觉理由：\*\* 涉及数据实体AuditTrail，该实体在系统中存在（基于Complaint记录），并连接其他实体（如跟踪修改），E-R图保持连接。\*\*完整引用理由：\*\* \*\*不完整引用\*\*，因为输入（“who accessed or modified”）未定义具体用户实体或权限细节；输出（audit trail）在SUPP-002提及，但存储需求（如CST-REG-002的12个月保留）是约束非功能需求，未在功能需求中定义输入输出。\*\*可测试理由：\*\* 输入（访问或修改操作）可模拟，输出（审计记录）可验证（如检查日志完整性和保留期），可测试性高。 |

| Enable Secure Communication (SUPP-003) | 否 | 是 | 否 | \*\*幻觉理由：\*\* 涉及数据实体Communication，该实体在系统中存在（用于投诉人通信），并连接Complaint实体，E-R图无孤立。\*\*完整引用理由：\*\* 输入（通信请求）和输出（安全渠道）在SUPP-003定义，外部约束（CST-IF-002的HTTPS/TLS 1.3）定义了安全标准，支持完整引用。\*\*可测试理由：\*\* \*\*不可测试\*\*，因为输入达成输出的转化不明确：需求仅描述“enable secure communication channels”，但未指定输入（如消息内容）如何转化为输出（安全传输），依赖外部因素（如网络配置）且无法直接验证功能；测试需要额外定义用例，但需求本身不足。 |

\*\*总结：\*\*

- 整个文档涉及的数据实体有5个：Complaint、Notification、AuditTrail、Department、Communication。

- 经过相同功能点合并后，功能需求共7条。

- 其中，幻觉需求有0条（所有需求均涉及可关联实体，E-R图弱连接）。

- 非幻觉需求有7条。

- 非幻觉需求中，完整引用的功能需求有5条（Notification Management和Maintain Audit Trail不完整引用）。

- 非幻觉需求中，可测试的功能需求有6条（Enable Secure Communication不可测试）。

理由补充：

- \*\*无幻觉需求\*\*：所有数据实体（如Notification、AuditTrail）均基于Complaint核心实体引入，并通过功能操作（如分类、通知）连接，E-R图始终弱连接。

- \*\*完整引用率不高\*\*：Notification Management和Maintain Audit Trail在输入输出细节上引用不足（如未定义通知格式或用户实体），符合“总可能有几条有问题”。

- \*\*可测试性良好\*\*：仅Enable Secure Communication不可测试，因其功能描述过于泛化，无法直接验证输入到输出转化。