# LLM 自动化评估结果

### 功能需求评估表

根据用户要求，我对系统功能需求文档进行了详细评估。评估过程包括以下步骤：

1. \*\*提取数据实体\*\*：从功能需求文档中识别系统内需要管理的数据实体（基于数据库表或核心对象），排除外部接口中的数据源或格式。共识别出 13 个数据实体：Customer、Administrator、Product、Category、Order、OrderDetail、PaymentInfo、ShoppingCart、CartItem、PluginInstance、EmailNotification、EventLog、Session。

2. \*\*合并功能需求\*\*：按照用户规则合并相同数据实体的相同功能点（例如，FR-07 和 FR-20 合并为“Checkout and Order Conversion”，FR-15 和 FR-21 合并为“Email Notification Sending”），并剔除系统通用设置类需求（无此类需求）。合并后，功能需求从原始 26 条减少到 24 条（合并两个功能对）。

3. \*\*识别幻觉需求\*\*：判断每个功能需求是否引入无法与现有数据实体关联的操作，导致 ER 图不连通（弱连接标准）。幻觉需求有 1 条。

4. \*\*完整引用评估\*\*：针对非幻觉功能点，判断是否引用未在需求规约中定义的输入、输出或功能（宽松标准，只要在文档中有提及即视为完整引用）。非幻觉需求中完整引用的有 22 条。

5. \*\*可测试性评估\*\*：针对非幻觉功能点，判断输入是否能转化为输出（宽松标准，允许外部依赖）。非幻觉需求中可测试的有 22 条。

6. \*\*严格评判\*\*：每个功能需求均独立评估，确保不遗漏。评估基于文档内容，稍严格处理边界情况（如 FR-16 被判定为幻觉，FR-17 和 FR-25 依赖外部 API 但视为可测试）。

评估表如下（功能需求基于合并后版本）：

| 功能需求 | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| FR-01: Customer Account Registration | 否 | 是 | 是 | 涉及 Customer 实体，输入输出在 DBMS 和 Plugin API 中定义，无未引用元素。 |

| FR-02: Customer Login | 否 | 是 | 是 | 涉及 Session 实体，输入输出在 Session Management 和 DBMS 中定义，测试可通过登录流程验证。 |

| FR-03: Customer Logout | 否 | 是 | 是 | 涉及 Session 实体，输入输出在 Session Management 中定义，测试可通过会话失效验证。 |

| FR-04: Customer Profile Management | 否 | 是 | 是 | 涉及 Customer 实体（更新/删除），输入输出在 DBMS 中定义，测试可通过数据变更验证。 |

| FR-05: Customer Purchase History View | 否 | 是 | 是 | 涉及 Order 和 OrderDetail 实体，输入输出在 DBMS 中定义，测试可通过查询历史订单验证。 |

| FR-06: Shopping Cart Management | 否 | 是 | 是 | 涉及 CartItem 实体，输入输出在 DBMS 中定义，测试可通过添加/移除商品验证。 |

| FR-07/20: Checkout and Order Conversion (合并) | 否 | 是 | 是 | 涉及 Order、OrderDetail 和 CartItem 实体，输入输出在 DBMS 和 Plugin API 中定义（FR-07 和 FR-20 重复，已合并），测试可通过订单创建验证。 |

| FR-08: Order Status View | 否 | 是 | 是 | 涉及 Order 实体，输入输出在 DBMS 中定义，测试可通过状态查询验证。 |

| FR-09: Order Cancellation | 否 | 是 | 是 | 涉及 Order 实体，输入输出在 DBMS 中定义，测试可通过状态更新验证。 |

| FR-10: Payment Information Management | 否 | 是 | 是 | 涉及 PaymentInfo 实体，输入输出在 DBMS 中定义，测试可通过支付数据管理验证。 |

| FR-12: Inventory Management | 否 | 是 | 是 | 涉及 Product 实体，输入输出在 DBMS 中定义，测试可通过库存操作验证。 |

| FR-13: Product Category Management | 否 | 是 | 是 | 涉及 Category 实体，输入输出在 DBMS 中定义，测试可通过分类操作验证。 |

| FR-14: Plugin Installation and Management | 否 | 是 | 是 | 涉及 PluginInstance 实体，输入输出在 DBMS 和 Plugin API 中定义，测试可通过插件安装验证。 |

| FR-15/21: Email Notification Sending (合并) | 否 | 是 | 是 | 涉及 EmailNotification 实体，输入输出在 Plugin API 和 Email Service 中定义（FR-15 和 FR-21 重复，已合并），测试可通过邮件发送模拟验证。 |

| FR-16: Discount Code Application | 否 | 是 | 是 | |

| FR-17: Plugin API Integration | 否 | 是 | 是 | 涉及 Plugin API，输入输出在软件接口中定义，无新实体，测试可通过 API 调用验证（依赖外部但可模拟）。 |

| FR-18: Session Management | 否 | 是 | 是 | 涉及 Session 实体，输入输出在 Session Management 模块中定义，测试可通过会话操作验证。 |

| FR-19: Order Detail Management | 否 | 是 | 是 | 涉及 OrderDetail 实体，输入输出在 DBMS 中定义，测试可通过详情更新验证。 |

| FR-22: Administrator Account Management | 否 | 是 | 是 | 涉及 Administrator 实体，输入输出在 DBMS 中定义，测试可通过账号操作验证。 |

| FR-23: System Event Logging | 否 | 是 | 是 | 涉及 EventLog 实体，输入输出在 DBMS 中定义，测试可通过日志记录验证。 |

| FR-24: Shopping Cart Creation and Access | 否 | 是 | 是 | 涉及 ShoppingCart 实体，输入输出在 DBMS 中定义，测试可通过购物车创建验证。 |

| FR-25: Product Availability Check | 否 | 是 | 是 | 涉及 Product 实体，输入输出在 Plugin API 中定义，测试可通过库存检查验证（依赖外部但可模拟）。 |

| FR-26: Shopping Cart Total Calculation | 否 | 是 | 是 | 涉及 CartItem 实体，输入输出在 DBMS 中定义，测试可通过总价计算验证。 |

| FR-27: Inventory Update on Purchase | 否 | 是 | 是 | 涉及 Product 实体，输入输出在 DBMS 中定义，测试可通过库存扣减验证。 |

\*\*总结\*\*：

- 整个文档涉及的数据实体有 \*\*13 个\*\*（Customer, Administrator, Product, Category, Order, OrderDetail, PaymentInfo, ShoppingCart, CartItem, PluginInstance, EmailNotification, EventLog, Session）。

- 经过相同功能点合并后，功能需求共 \*\*24 条\*\*（原始 26 条，合并 FR-07 与 FR-20、FR-15 与 FR-21）。

- 其中，幻觉需求有 \*\*0 条\*\*（FR-16）。

- 非幻觉需求有 \*\*24 条\*\*。

- 非幻觉需求中完整引用的功能需求有 \*\*24 条\*\*（FR-17 和 FR-25 依赖外部接口但文档中有定义，视为完整引用；所有非幻觉均完整引用）。

- 非幻觉需求中可测试的功能需求有 \*\*24 条\*\*（所有非幻觉均可测试；FR-15/21 和 FR-25 依赖外部 API 但可模拟测试）。

\*\*评估说明\*\*：

- 幻觉需求仅 FR-16，因引入未定义的折扣实体，破坏 ER 图连通性。

- 非幻觉需求中，FR-17 和 FR-25 完整引用和可测试，因 Plugin API 在外部接口中明确定义。

- 可测试性评估宽松处理：外部依赖（如 API）视为可测试。

- 合并规则严格：仅合并重复操作（如订单转换和邮件发送），其他 CRUD 未合并避免过度泛化。