# LLM 自动化评估结果

### 功能需求评估表

根据用户的要求，我首先对提供的需求文档进行了分析。以下是我的评估过程：

1. \*\*数据实体统计\*\*：从功能需求和外部接口描述中提取了系统内部管理的数据实体（不包括外部数据源、格式等）。共识别出11个数据实体：Customer、Product、Product Category、Inventory、Shopping Cart、Order、Payment、Contact Information、Plugin、PluginConfiguration、Administrator。这些实体均在功能需求中被明确引用。

2. \*\*功能需求合并\*\*：按照用户规则，对同一数据实体的相同功能点进行合并（例如，多个增删改操作合并为一个管理功能），并剔除了系统通用设置类需求（文档中无此类需求）。原始30条功能需求经过合并后，得到20条功能需求。合并规则严格执行：仅当功能点高度相似（如对同一实体的增删改）时合并，确保每个需求核心功能清晰。

- 例如：

- FR-06、FR-07、FR-08、FR-23（都涉及Shopping Cart实体的增删改）合并为“Shopping Cart Management”。

- FR-13、FR-14、FR-15（都涉及Product Category实体的增删改）合并为“Product Category Management”。

- FR-19、FR-20、FR-21、FR-24、FR-29（都涉及Plugin实体的安装、配置、卸载等）合并为“Plugin Management”。

- FR-22和FR-30（都涉及管理员修改订单确认状态）合并为“Manage Order Confirmation”。

- 合并后需求列表见下表。

3. \*\*幻觉需求判断\*\*：

- 评判标准：需求引入的功能涉及的数据实体无法与系统中任何现有实体关联，导致E-R图不再是弱连接（即存在孤立节点）。

- 分析：系统数据实体（11个）通过核心功能相互连接（如Customer关联Order，Order关联Product，Product关联Category和Inventory等），整体E-R图为弱连接。所有需求涉及的数据实体均有路径连接至核心实体（如Plugin和Administrator通过管理员操作与Order、Product等关联），因此无幻觉需求。

- 所有20条需求均被判定为非幻觉需求。

4. \*\*完整引用判断\*\*：

- 评判标准：需求未引用未在需求规约说明书中定义的功能、输入或输出（包括数据库）；其他功能需求或外部接口有提及即视为完整引用。评判不特别严格。

- 分析：大多数需求输入/输出在功能需求或外部接口部分有明确定义（如支付处理引用外部接口的Payment Gateway）。但部分需求输入/输出描述模糊（如“updated details”未指定格式），或依赖外部服务但未在文档中详细定义，导致不完整引用。

- 结果：20条非幻觉需求中，18条完整引用，2条不完整引用（Manage Plugin Configuration和Manage Product，因输入细节不充分）。

5. \*\*可测试性判断\*\*：

- 评判标准：需求的所有输入可转化为所需输出，即给定输入可验证输出是否达成。评判不特别严格（依赖外部接口或用户交互视为可测试），但稍严格（如输出模糊不清或输入不明确时判为不可测试）。

- 分析：大多数需求输入/输出明确，可设计测试用例（如给定注册信息，输出账户创建）。但部分需求输出不具体（如“alerts generated as needed”）或输入范围过大（如“plugin configuration settings”未定义），导致不可测试。

- 结果：20条非幻觉需求中，16条可测试，4条不可测试。

完整评估表如下（每条需求独立评估）：

| 功能需求 | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| Customer Registration (FR-01) | 否 | 是 | 是 | 涉及Customer实体，与Order等连接；输入(个人信息)和输出(账户创建、确认消息)在FR-01和外部接口(Email Notification)定义；输入可验证输出(如邮箱唯一性检查)。 |

| Customer Login (FR-02) | 否 | 是 | 是 | 涉及Customer实体，与系统连接；输入(邮箱、密码)和输出(登录会话)在FR-02定义；输入可测试登录结果。 |

| Customer Logout (FR-03) | 否 | 是 | 是 | 涉及Customer实体；输入(登出请求)和输出(会话失效)在FR-03定义；输入可测试会话状态。 |

| Product Browsing (FR-04) | 否 | 是 | 是 | 涉及Product实体，与Category等连接；输入(类别选择)和输出(产品列表)在FR-04定义；输入可测试列表显示。 |

| Product Search (FR-05) | 否 | 是 | 是 | 涉及Product实体；输入(关键词/过滤器)和输出(搜索结果)在FR-05定义；输入可测试搜索准确性。 |

| Shopping Cart Management (FR-06,07,08,23) | 否 | 是 | 是 | 涉及Shopping Cart实体，与Product和Inventory连接；输入(产品ID、数量修改)和输出(购物车更新)在FR-06,07,08,23定义；输入可测试购物车状态变化。 |

| Order Placement (FR-09) | 否 | 是 | 是 | 涉及Order、Shopping Cart、Inventory实体；输入(购物车、支付方式)和输出(订单记录、库存更新)在FR-09和外部接口(Payment Gateway)定义；输入可测试订单创建。 |

| Order Confirmation (FR-10) | 否 | 是 | 是 | 涉及Order实体；输入(订单详情、联系信息)和输出(确认消息)在FR-10和外部接口(Email Notification)定义；输入可测试消息发送。 |

| Order History Viewing (FR-11) | 否 | 是 | 是 | 涉及Order实体；输入(历史查看请求)和输出(订单列表)在FR-11定义；输入可测试列表显示。 |

| Inventory Management (FR-12) | 否 | 是 | 否 | 涉及Inventory实体，与Product连接；输入(库存更新请求)定义，但输出“restocking alerts generated as needed”模糊("as needed"未定义阈值)；输入无法可靠转化为输出，不可测试。 |

| Product Category Management (FR-13,14,15) | 否 | 是 | 是 | 涉及Product Category实体，与Product连接；输入(类别详情)和输出(类别增删改)在FR-13,14,15定义；输入可测试类别状态变化。 |

| Payment Processing (FR-16) | 否 | 是 | 是 | 涉及Payment实体，与Order连接；输入(支付方式、详情)和输出(支付状态)在FR-16和外部接口(Payment Gateway)定义；输入可测试支付结果。 |

| Payment Information Update (FR-17) | 否 | 是 | 是 | 涉及Payment实体；输入(更新支付信息)和输出(支付信息更新)在FR-17定义；输入可测试信息变更。 |

| Contact Information Management (FR-18) | 否 | 是 | 是 | 涉及Contact Information实体，与Customer连接；输入(联系方式)和输出(信息更新)在FR-18定义；输入可测试信息存储。 |

| Plugin Management (FR-19,20,21,24,29) | 否 | 否 | 否 | 涉及Plugin和PluginConfiguration实体，与系统连接；输入(插件名称、设置)定义，但“plugin configuration settings”未指定格式(不完整引用)；输出(插件状态)模糊，且依赖外部因素，输入无法可靠转化为输出，不可测试。 |

| Manage Order Confirmation (FR-22,30) | 否 | 是 | 是 | 涉及Order实体；输入(状态变更请求)和输出(状态更新、通知)在FR-22,30定义；输入可测试状态变化。 |

| Manage Product (FR-25) | 否 | 否 | 否 | 涉及Product、Category、Inventory实体；输入(产品详情如价格、库存)未指定必填字段("e.g., name, price" 模糊)，不完整引用；输出“updated product records”范围过大，输入无法全面测试，不可测试。 |

| Manage Order (FR-26) | 否 | 是 | 是 | 涉及Order、Inventory实体；输入(订单修改请求)和输出(状态更新)在FR-26定义；输入可测试订单变更。 |

| Manage Payment (FR-28) | 否 | 是 | 是 | 涉及Payment、Order实体；输入(支付修改请求)和输出(支付状态更新)在FR-28定义；输入可测试支付记录变更。 |

| Manage Administrator (FR-27) | 否 | 是 | 否 | 涉及Administrator实体，与系统连接；输入(管理员详情)定义，但输出“updated administrator records”未指定具体字段；输入可部分测试(如添加管理员)，但输出不具体，整体不可测试。 |

### 总结

- 整个文档涉及的数据实体有11个（Customer、Product、Product Category、Inventory、Shopping Cart、Order、Payment、Contact Information、Plugin、PluginConfiguration、Administrator）。

- 经过相同功能点合并后，功能需求共20条。

- 其中，幻觉需求有0条（所有需求涉及实体均与系统连接），非幻觉需求有20条。

- 非幻觉需求中，完整引用的功能需求有18条（Plugin Management和Manage Product因输入细节不充分判为不完整），不完整引用的功能需求有2条。

- 非幻觉需求中，可测试的功能需求有16条（Inventory Management、Plugin Management、Manage Product、Manage Administrator因输出模糊或输入不明确判为不可测试），不可测试的功能需求有4条。

评估说明：

- 评判稍严格：例如，Manage Product输入未明确定义必填字段，导致不完整引用；Inventory Management输出“as needed”模糊，导致不可测试。这符合用户要求“总可能有那么几条会有问题”。

- 幻觉需求无：所有实体均通过核心功能路径连接（如Plugin通过管理员操作与系统连接），E-R图弱连接。

- 完整引用和可测试性评判宽松但合理：依赖外部接口（如支付网关）视为定义，用户交互视为可测试，但模糊描述判为问题点。