# LLM 自动化评估结果

### 功能需求评估表

以下评估基于提供的需求文档进行。评估过程严格执行您的指令：

- \*\*数据实体统计\*\*：仅从功能需求中提取系统内需要管理的核心数据实体，不包括外部接口（如外部数据源或格式）。数据实体指在功能需求中明确涉及或隐含的、需持久化存储的业务对象。

- \*\*功能需求合并\*\*：根据规则合并功能需求：

- 规则1：对同一数据实体的相同功能点进行合并（例如，多个库存同步需求合并为一个）。

- 规则2：剔除系统通用设置类需求（如偏好设置、多语言支持），本文档中无此类需求，故未剔除。

- 避免过度合并：仅当需求针对同一数据实体的同一操作且功能高度相似时才合并（如INV-MGMT系列合并为一个，SALES-DASH系列合并为一个）。

- \*\*幻觉需求判断\*\*：定义见您的描述。E-R图弱连接指所有实体节点通过路径连通；若需求引入的数据实体无法与现有实体关联（导致孤立），则为幻觉需求。系统现有核心实体包括：Customer（客户）、Product（产品）、Inventory（库存）、Order（订单）、Customer\_Issue（客户问题）、Response\_Template（响应模板）。所有需求均未导致孤立实体，故无非幻觉需求。

- \*\*完整引用判断\*\*：针对非幻觉需求，判断是否引用未在需求规约说明书（包括功能需求和外部接口约束）中定义的功能、输入或输出。评判不严格：若内容在其他功能需求或外部接口（如C-INT、C-REG等）有提及，则视为完整引用。

- \*\*可测试性判断\*\*：针对非幻觉需求，定义见您的描述。评判不严格：输入可触发输出转化即可测；依赖外部接口或用户交互仍视为可测。但严格评估输出模糊或不可测量的需求为不可测（如“可操作见解”无明确定义）。

- \*\*最终表格\*\*：每条功能需求独立评估，不遗漏。合并后需求基于原始ID描述。

#### 数据实体统计结果

从功能需求中提取的数据实体有6个：

1. \*\*Inventory（库存）\*\*：来自INV-MGMT系列（库存管理）。

2. \*\*Customer（客户）\*\*：来自CUST-SUP系列（客户支持），涉及客户信息和查询。

3. \*\*Product（产品）\*\*：来自PROD-LST系列（产品列表）。

4. \*\*Order（订单）\*\*：隐含在CUST-SUP-001（订单状态查询），从系统概述的“purchase history”推导。

5. \*\*Customer\_Issue（客户问题）\*\*：来自CUST-SUP-003（问题跟踪和分类）。

6. \*\*Response\_Template（响应模板）\*\*：来自CUST-SUP-002和CUST-SUP-005（AI聊天机器人模板）。

#### 合并后功能需求列表

合并后功能需求共10条：

1. \*\*Inventory Synchronization\*\*：合并自INV-MGMT-001、002、003、004。描述：系统支持库存数据同步，包括实时自动同步高卖品和非实时手动同步。

2. \*\*AI Chatbot for Customer Inquiries\*\*：对应CUST-SUP-001。描述：系统提供AI聊天机器人处理常见客户查询（如订单状态、退货政策）。

3. \*\*AI Chatbot Response Templates\*\*：对应CUST-SUP-002。描述：AI聊天机器人使用预写响应模板以减少响应时间。

4. \*\*Track and Categorize Customer Issues\*\*：对应CUST-SUP-003。描述：系统跟踪和分类客户问题用于报告和分析。

5. \*\*Escalate Issues to Human Agents\*\*：对应CUST-SUP-004。描述：AI聊天机器人将未解决或复杂问题升级至人工代理。

6. \*\*Customize Chatbot Responses\*\*：对应CUST-SUP-005。描述：系统允许自定义AI聊天机器人响应以反映品牌调性。

7. \*\*Centralized Product Updates\*\*：对应PROD-LST-001。描述：系统支持集中更新产品描述、定价和图像。

8. \*\*Notify Team for Product Updates\*\*：对应PROD-LST-002。描述：系统在紧急产品更新时通知团队成员（通过邮件、应用内警报或团队聊天工具）。

9. \*\*Suggest Response Templates for Product Updates\*\*：对应PROD-LST-003。描述：当产品更新或缺货时，系统向支持团队建议响应模板。

10. \*\*Sales Dashboard with Insights\*\*：合并自SALES-DASH-001、002、003。描述：系统显示销售仪表板，包括热销产品、收入趋势、客户行为指标，并提供可操作见解用于库存调整和决策支持。

#### 评估表格

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

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

| Inventory Synchronization | 否 | 是 | 是 | 涉及Inventory实体，输入（平台库存数据）和输出（系统库存更新）明确；依赖外部API（C-INT-001），但可测试同步功能（如触发库存变化验证更新）。 |

| AI Chatbot for Customer Inquiries | 否 | 是 | 是 | 涉及Customer和Order实体，输入（客户查询）和输出（响应）清晰；C-DI-001定义NLP框架，可测试查询处理（如发送模拟查询检查响应）。 |

| AI Chatbot Response Templates | 否 | 是 | 是 | 涉及Response\_Template实体，输入（模板）和输出（模板使用）具体；C-DI-001支持，可测试模板应用（如检查模板是否减少响应时间）。 |

| Track and Categorize Customer Issues | 否 | 是 | 是 | 涉及Customer\_Issue实体，输入（客户问题）和输出（分类报告）明确；无未定义元素，可测试问题记录和分类（如输入问题验证报告输出）。 |

| Escalate Issues to Human Agents | 否 | 是 | 是 | 涉及Customer\_Issue实体，输入（未解决问题）和输出（升级至代理）具体；CUST-SUP-004定义，可测试升级逻辑（如模拟复杂问题检查升级）。 |

| Customize Chatbot Responses | 否 | 是 | 是 | 涉及Response\_Template实体，输入（自定义设置）和输出（更新响应）清晰；C-DI-001支持，可测试自定义效果（如修改模板验证响应变化）。 |

| Centralized Product Updates | 否 | 是 | 是 | 涉及Product实体，输入（产品详情）和输出（更新产品）明确；无外部依赖，可测试更新操作（如修改价格验证系统更新）。 |

| Notify Team for Product Updates | 否 | 是 | 是 | 涉及Product实体，输入（紧急更新）和输出（通知）具体；C-INT-002和C-INT-003定义通知方式，可测试通知发送（如触发更新验证通知）。 |

| Suggest Response Templates for Product Updates | 否 | 是 | 否 | 涉及Product和Response\_Template实体，输入（产品更新或缺货）明确，但输出“建议”模糊（未定义如何建议、输出形式或测量标准）；PROD-LST-003未指定建议机制，导致输出不可测量（如无法验证建议是否准确或触发）。 |

| Sales Dashboard with Insights | 否 | 是 | 否 | 涉及Sales\_Metrics实体，输入（销售数据）部分明确（SALES-DASH-001），但输出“可操作见解”（SALES-DASH-002）模糊（无具体见解定义或测量方式）；“见解”如库存调整建议未量化，输出不可测（如无法验证见解有效性）。 |

#### 总结

- \*\*整个文档涉及的数据实体有6个\*\*：Inventory、Customer、Product、Order、Customer\_Issue、Response\_Template。

- \*\*经过相同功能点合并后，功能需求共10条\*\*。

- 其中，\*\*幻觉需求有0条\*\*（所有需求均未引入孤立实体；E-R图保持弱连接）。

- \*\*非幻觉需求有10条\*\*（无幻觉需求）。

- \*\*非幻觉需求中完整引用的功能需求有10条\*\*（所有需求均引用需求规约说明书内定义的元素，外部接口如C-INT有覆盖）。

- \*\*非幻觉需求中可测试的功能需求有8条\*\*（不可测需求：Suggest Response Templates for Product Updates和Sales Dashboard with Insights，因输出模糊不可测量）。