# LLM 自动化评估结果

### 功能需求评估表

以下评估基于用户提供的系统概述和需求文档进行。评估过程严格遵守您的指示：

- \*\*数据实体识别\*\*：从功能需求中提取系统内管理的核心数据实体（不包括外部接口）。识别出的数据实体包括：用户数据（Customer）、订单数据（Order）、产品数据（Product）、库存数据（Inventory）、日志数据（Log）、查询数据（Inquiry）。购物车（Cart）在需求中未作为独立实体提及（作为订单的一部分处理），支付信息（Payment Information）嵌入用户数据或订单数据。因此，数据实体共6个。

- \*\*功能需求合并\*\*：依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合并（如1.1和1.2均涉及敏感数据加密/解密，合并为一个需求）。

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

- 合并后功能需求共9条（原始需求10条，其中1.1和1.2合并为FR1）。合并时确保不过度合并，每个需求清晰表达核心功能。

- \*\*幻觉需求判断\*\*：标准为需求引入的功能操作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图非弱连接。现有数据实体（Customer、Order、Product、Inventory、Log、Inquiry）覆盖所有需求功能点，因此无幻觉需求。

- \*\*完整引用判断\*\*：标准为需求未引用未在需求规约说明书（给定文档）中定义的功能、输入或输出（包括数据库）。评判不严格，只要其他功能需求或约束（外部接口）提到即视为完整引用。

- \*\*可测试性判断\*\*：标准为需求的所有输入可达成所需输出的转化。评判不严格，依赖于外部接口或用户交互的也视为可测试。

- \*\*评估严格性\*\*：依据您的指示，我进行了稍严格评判：FR5（消息路由）和FR8（通信渠道）在完整引用和可测试性上存在潜在问题（理由详见表）。其他需求均符合要求。

#### 合并后功能需求列表及评估

| 功能需求编号 | 功能需求描述 (合并后) | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| FR1 | 系统应使用AES-256加密和解密敏感数据，并保持数据完整性（合并自1.1和1.2）。 | 否 | 是 | 是 | \*\*幻觉需求：\*\* 涉及敏感数据（如用户数据、支付信息），属于用户数据（Customer）实体的一部分，与现有实体关联。\*\*完整引用：\*\* 输入（敏感数据）和输出（加密/解密数据）在系统概述中定义（用户信息、支付信息）。\*\*可测试：\*\* 输入数据可加密/解密，输出可验证（如测试加密强度）。 |

| FR2 | 系统应提供审计日志，记录用户操作和系统事件（来自1.3）。 | 否 | 是 | 是 | \*\*幻觉需求：\*\* 涉及日志数据（Log）实体，与系统概述中的日志功能关联。\*\*完整引用：\*\* 输入（用户操作、系统事件）和输出（日志）在需求中明确定义。\*\*可测试：\*\* 输入操作可触发，日志输出可检查（如模拟用户操作后验证日志条目）。 |

| FR3 | 系统应允许用户通过专用界面请求访问或删除其数据（来自1.4）。 | 否 | 是 | 是 | \*\*幻觉需求：\*\* 涉及用户数据（Customer）实体，直接关联现有实体。\*\*完整引用：\*\* 输入（用户请求）和输出（数据访问/删除）在需求中定义，未引用未定义元素。\*\*可测试：\*\* 输入请求可模拟，输出数据变更可验证（如检查数据删除效果）。 |

| FR4 | 系统应在收到库存API消息后1秒内更新实时订单状态（来自2.1）。 | 否 | 是 | 是 | \*\*幻觉需求：\*\* 涉及订单数据（Order）实体，与系统概述中的订单管理关联。\*\*完整引用：\*\* 输入（库存API消息）在约束C-INT-001中提及，输出（订单状态更新）在需求中定义。\*\*可测试：\*\* 输入API消息可模拟，输出状态更新可计时验证（如发送测试消息检查响应时间）。 |

| FR5 | 系统应使用定义协议聚合和路由用户界面与外部服务之间的消息（来自2.2）。 | 否 | 否 | 否 | \*\*幻觉需求：\*\* 涉及消息路由功能，但消息可关联到订单（Order）或查询（Inquiry）实体（如订单状态消息），未引入新孤立实体。\*\*完整引用：\*\* 输入/输出（消息）未在文档中明确定义；协议仅在需求中提为“defined”但未详述，且约束C-INT-003提到消息队列但不直接定义协议，因此不完整引用。\*\*可测试：\*\* 输入消息和协议未明确定义，输出路由无法直接验证（如协议细节缺失，测试依赖未定义的输入）。 |

| FR6 | 系统应提供常见客户查询列表及预定义答案（来自3.1）。 | 否 | 是 | 是 | \*\*幻觉需求：\*\* 涉及查询数据（Inquiry）实体，与系统概述中的客户查询功能关联。\*\*完整引用：\*\* 输入（无明确输入）和输出（查询列表）在需求中定义，未引用未定义元素。\*\*可测试：\*\* 输出列表可检查存在性和内容（如验证列表是否包含预定义答案）。 |

| FR7 | 系统应允许客服代表手动撰写响应客户查询（邮件或聊天），成功率至少99%（来自3.2）。 | 否 | 是 | 是 | \*\*幻觉需求：\*\* 涉及查询数据（Inquiry）实体，响应作为查询的一部分关联。\*\*完整引用：\*\* 输入（客服输入）和输出（响应）在需求中定义，成功率指标明确。\*\*可测试：\*\* 输入查询和撰写操作可模拟，输出响应可统计成功率（如测试多场景计算成功比例）。 |

| FR8 | 系统应基于查询类型确定适当通信渠道，标准在配置设置中定义（来自3.3）。 | 否 | 是 | 否 | \*\*幻觉需求：\*\* 涉及查询数据（Inquiry）实体，渠道作为属性关联。\*\*完整引用：\*\* 输入（查询类型）和输出（渠道）在需求中定义，配置标准在需求中提及（尽管未详述）。\*\*可测试：\*\* 输入查询类型可设置，但输出渠道依赖于未定义的配置设置（如配置标准未在文档中说明），无法完全验证转化。 |

| FR9 | 系统应在库存水平达到临界点时启动重新订购流程，通过采购请求到库存API，成功率至少99%（来自4.1）。 | 否 | 是 | 是 | \*\*幻觉需求：\*\* 涉及库存数据（Inventory）实体，与系统概述中的库存管理关联。\*\*完整引用：\*\* 输入（库存水平）和输出（采购请求）在需求中定义，库存API在约束C-INT-002和C-INT-004提及。\*\*可测试：\*\* 输入库存水平可设置，输出请求可验证成功率和流程（如模拟临界库存检查API调用）。 |

#### 总结

- \*\*整个文档涉及的数据实体数量\*\*：6个（Customer、Order、Product、Inventory、Log、Inquiry）。

- \*\*经过相同功能点合并后，功能需求共\*\*：9条。

- \*\*其中，幻觉需求有\*\*：0条（所有需求均关联现有实体，E-R图保持弱连接）。

- \*\*非幻觉需求有\*\*：9条（幻觉需求为0，故全部为非幻觉）。

- \*\*非幻觉需求中完整引用的功能需求有\*\*：8条（FR5不完整引用，因协议未定义）。

- \*\*非幻觉需求中可测试的功能需求有\*\*：7条（FR5和FR8不可测试；FR5因输入/输出未明确定义，FR8因依赖未定义配置）。

\*\*评估说明\*\*：

- 评估稍严格：FR5（消息路由）在完整引用和可测试性上判为否，因协议未定义导致输入/输出不明确；FR8（通信渠道）在可测试性上判为否，因配置标准未详述。其他需求均符合标准。

- 无系统通用设置类需求被统计或误判为幻觉。

- 所有功能需求均被评估，无遗漏。