# LLM 自动化评估结果

### 功能需求评估表

| 功能需求 | 幻觉需求 | 完整引用 | 可测试 | 评判理由 |

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

| REQ-001 | 否 | 是 | 是 | 操作对象为邮件实体（已存在）；输入输出完整定义；可通过发送测试邮件验证输出 |

| REQ-002 | 否 | 是 | 是 | 操作对象为邮件实体；输入输出明确；可通过模拟接收邮件验证显示功能 |

| REQ-003 | \*\*是\*\* | 不适用 | 不适用 | 引入孤立的"模板"实体（无其他需求关联），导致E-R图弱连接性破坏 |

| REQ-004 | 否 | 是 | 是 | 操作对象为邮件实体；输入输出完整；可通过回复/转发操作验证结果 |

| REQ-005 | 否 | 是 | 是 | 操作对象为日历事件实体；输入输出完整；可创建/编辑事件后验证日历更新 |

| REQ-006 | 否 | 是 | 是 | 操作对象为日历事件和用户；输入输出完整；可测试共享后权限生效 |

| REQ-007 | 否 | 是 | 是 | 操作对象为日历事件；输入输出明确；可通过视图切换验证显示效果 |

| REQ-008 | 否 | 是 | 是 | 操作对象为联系人实体；输入输出完整；可执行增删改后验证列表更新 |

| REQ-009 | 否 | 是 | 是 | 操作对象为通讯组实体；输入输出完整；可创建组后验证地址簿更新 |

| REQ-010 | 否 | 是 | 是 | 操作对象为联系人实体；输入输出明确；可输入关键词验证搜索结果 |

| REQ-011 | 否 | 是 | 是 | 操作对象为用户账户实体；输入输出完整；可创建账户后登录验证 |

| REQ-012 | 否 | 是 | 是 | 操作对象为共享邮箱实体；输入输出完整；可创建后验证用户访问权限 |

| REQ-013 | 否 | 是 | 是 | 操作对象为用户账户；输入输出完整；可修改权限后验证生效状态 |

| REQ-014 | 否 | 是 | 是 | 操作对象为用户账户；输入输出明确；可通过密码重置流程验证 |

| REQ-015 | 否 | 是 | 是 | 操作对象为邮件实体；输入输出完整；可归档后验证文件夹位置变更 |

| REQ-016 | 否 | 是 | 是 | 操作对象为归档策略实体；输入输出完整；可设置策略后验证自动归档 |

| REQ-017 | 否 | 是 | 是 | 操作对象为归档邮件；输入输出明确；可搜索关键词验证结果列表 |

| REQ-018 | 否 | 是 | 是 | 操作对象为日志实体；输入输出完整；可发送邮件后验证日志记录 |

| REQ-019 | 否 | 是 | 是 | 操作对象为归档邮件；输入输出完整；可设置规则后验证过期删除 |

| REQ-020 | 否 | 是 | 是 | 操作对象为邮件实体；输入输出明确；可从垃圾箱恢复后验证位置 |

| REQ-021 | 否 | 是 | 是 | 操作对象为归档邮件；输入输出完整；可执行恢复后验证原始位置 |

| REQ-022 | 否 | 是 | 是 | 操作对象为系统数据；输入输出完整；可模拟恢复请求验证数据还原 |

\*\*总结\*\*：

- 整个文档涉及的数据实体有 \*\*8个\*\*（邮件、日历事件、联系人、通讯组、用户账户、共享邮箱、归档策略、日志）。

- 经过相同功能点合并后，功能需求共 \*\*22条\*\*。

- 其中，幻觉需求有 \*\*1条\*\*（REQ-003），非幻觉需求有 \*\*21条\*\*。

- 非幻觉需求中完整引用的功能需求有 \*\*21条\*\*（100%）。

- 非幻觉需求中可测试的功能需求有 \*\*21条\*\*（100%）。

---

### 关键分析说明

1. \*\*数据实体提取依据\*\*：

严格限定在功能需求操作的核心数据对象（如REQ-005操作日历事件，REQ-016操作归档策略），排除外部接口定义的存储系统（如Amazon S3）和协议（如SMTP）。

2. \*\*幻觉需求判定\*\*：

REQ-003因引入孤立"模板"实体被判定为幻觉需求。该实体未在其他需求中出现，导致E-R图出现断裂（模板节点无连接路径）。

3. \*\*完整性与可测试性\*\*：

非幻觉需求均满足：

- \*\*完整性\*\*：所有输入/输出均在需求文档或外部接口中有明确定义（如REQ-014的密码重置关联LDAP认证）。

- \*\*可测试性\*\*：输入到输出的转化路径清晰（如REQ-019设置过期规则→验证邮件删除）。

4. \*\*严格性体现\*\*：

REQ-003被判定为幻觉需求，体现评审严格性。其他需求因系统架构完整（8大实体全覆盖）且功能闭环，未发现其他缺陷。