# LLM 自动化评估结果

### 数据实体统计

1. \*\*Email\*\*（邮件）

2. \*\*Event\*\*（事件）

3. \*\*Task\*\*（任务）

4. \*\*Folder/Label\*\*（文件夹/标签）

5. \*\*Category\*\*（分类）

6. \*\*Custom Phrase\*\*（自定义短语）

7. \*\*Configuration\*\*（配置）

### 功能需求合并与归类

1. \*\*FR-1, FR-2, FR-3, FR-4, FR-5\*\*：合并为 \*\*"事件检测功能（含自定义短语管理）"\*\*（涉及数据实体：\*\*Email, Event, Custom Phrase\*\*）

2. \*\*FR-6, FR-7\*\*：合并为 \*\*"任务和事件管理（含统一视图）"\*\*（涉及数据实体：\*\*Task, Event\*\*）

3. \*\*FR-8, FR-9\*\*：合并为 \*\*"邮件自动分类与手动分类"\*\*（涉及数据实体：\*\*Email, Folder/Label\*\*）

4. \*\*FR-10\*\*：保留为 \*\*"邮件分类标签管理"\*\*（涉及数据实体：\*\*Email, Category\*\*）

5. \*\*FR-11, FR-12\*\*：合并为 \*\*"配置模块（含规则管理与导入导出）"\*\*（涉及数据实体：\*\*Configuration\*\*）

剔除通用设置类需求：无（文档中无多语言、偏好设置等无关需求）。

---

### 幻觉需求判断

| 功能需求 | 幻觉需求 | 判断理由 |

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

| 事件检测功能（含自定义短语管理） | 否 | 数据实体（Email, Event, Custom Phrase）均与系统核心功能直接关联。 |

| 任务和事件管理（含统一视图） | 否 | 数据实体（Task, Event）与日历和邮件管理紧密相关。 |

| 邮件自动分类与手动分类 | 否 | 数据实体（Email, Folder/Label）是邮件系统的核心组成部分。 |

| 邮件分类标签管理 | 否 | 数据实体（Email, Category）扩展了邮件管理功能，与系统关联明确。 |

| 配置模块（含规则管理与导入导出） | 否 | 数据实体（Configuration）是系统配置的基础，与其他功能直接交互。 |

\*\*总结\*\*：无幻觉需求。

---

### 完整引用判断

| 功能需求 | 完整引用 | 判断理由 |

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

| 事件检测功能（含自定义短语管理） | 是 | 所有输入（邮件内容、自定义短语）和输出（事件提取）均在需求中明确定义。 |

| 任务和事件管理（含统一视图） | 是 | 功能依赖的实体（Task, Event）和操作（增删改查）已完整描述。 |

| 邮件自动分类与手动分类 | 是 | 输入（邮件、规则）和输出（文件夹分配）均明确，且无外部未定义依赖。 |

| 邮件分类标签管理 | 是 | 预定义和自定义分类均在需求中说明，操作边界清晰。 |

| 配置模块（含规则管理与导入导出） | 是 | JSON格式和导入导出功能已明确，无隐藏依赖。 |

\*\*总结\*\*：所有需求均为完整引用。

---

### 可测试性判断

| 功能需求 | 可测试 | 判断理由 |

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

| 事件检测功能（含自定义短语管理） | 是 | 可通过模拟邮件输入验证事件提取和自定义短语的匹配逻辑。 |

| 任务和事件管理（含统一视图） | 是 | 增删改查操作和视图排序可通过用户交互测试。 |

| 邮件自动分类与手动分类 | 是 | 自动分类响应时间（≤2秒）和手动分配功能可实测。 |

| 邮件分类标签管理 | 是 | 预定义标签和自定义标签的分配功能可验证。 |

| 配置模块（含规则管理与导入导出） | 是 | JSON导入导出及规则配置可通过文件操作测试。 |

\*\*总结\*\*：所有需求均可测试。

---

### 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 |

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

| 事件检测功能（含自定义短语管理） | 否 | 是 | 是 | 核心功能，输入输出明确，可模拟测试。 |

| 任务和事件管理（含统一视图） | 否 | 是 | 是 | 实体与操作完整，交互可验证。 |

| 邮件自动分类与手动分类 | 否 | 是 | 是 | 分类逻辑和响应时间可量化测试。 |

| 邮件分类标签管理 | 否 | 是 | 是 | 标签分配功能边界清晰。 |

| 配置模块（含规则管理与导入导出） | 否 | 是 | 是 | 配置操作和文件处理可验证。 |

\*\*总结\*\*：

- 合并后功能需求共 \*\*5条\*\*。

- 幻觉需求 \*\*0条\*\*，非幻觉需求 \*\*5条\*\*。

- 非幻觉需求中完整引用 \*\*5条\*\*，且全部可测试。

- 涉及数据实体 \*\*7个\*\*（Email, Event, Task, Folder/Label, Category, Custom Phrase, Configuration）。